

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 13 November, 2006	3. REPORT TYPE AND DATES COVERED Final Progress Report (June 1, 1998 - May 31, 2004)		
4. TITLE AND SUBTITLE Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures		5. FUNDING NUMBERS Grant No.: DAAG55-98-1-0293		
6. AUTHORS Armand R. Tanguay, Jr., B. Keith Jenkins, Christoph von der Malsburg, Bartlett Mel, Irving Biederman, John O'Brien, Anupam Madhukar				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Southern California 920 W. 37th Street, Seaver Science Center 520 University Park, MC-0483 Los Angeles, CA 90089-0483		8. PERFORMING ORGANIZATION REPORT NUMBER USC-OMDL-01-MURI-Final		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211		10. SPONSORING / MONITORING AGENCY REPORT NUMBER  38565.1-EL-MUR		
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be considered as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT  Approved for public release; distribution unlimited		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words) The goal of this research program was to develop novel algorithms, architectures, and hardware for a truly smart camera, with inherent capability for semi-autonomous object recognition as well as optimal image capture. In this research, we combined striking advances in the understanding of the mechanisms of biological vision systems with similar advances in hybrid electronic/photonic packaging technology, in order to develop adaptive, artificial, biologically-inspired vision systems. A key research program objective, therefore, was to establish and address the fundamental scientific and technological issues that currently inhibit the implementation of such adaptive optoelectronic eyes. Several novel approaches to the vertical integration of multiple silicon VLSI chips into a hybrid electronic/photonic multichip module by means of dense 3-D photonic interconnections were pursued. In this approach, local and quasi-local connectivity between layers is accomplished by using diffractive optical structures that provide for both point-to-point interconnections and weighted fan-out within a local neighborhood. During this research program, significant progress was achieved in the definition of scientific and technological hurdles; the establishment of commonalities among several low-level, intermediate-level, and high-level vision models; the mapping of key functionalities onto the emerging hybrid electronic/photonic architecture; the characterization of key components such as vertical cavity surface-emitting laser arrays (VCSELs) and diffractive optical elements (DOEs); and key steps in the integration of hybrid electronic/photonic multichip modules that are capable of implementing these vision models.				
14. SUBJECT TERMS Adaptive optoelectronic eyes; smart cameras; hybrid electronic/photonic multichip modules; vision systems; diffractive optical elements; DOEs; MCMs; modulator arrays; laser arrays; electronic/photonic interconnections; vertical cavity surface emitting laser arrays; VCSEL arrays			15. NUMBER OF PAGES 125	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

**Adaptive Optoelectronic Eyes:  
Hybrid Sensor/Processor Architectures**

**University of Southern California  
(Principal Investigator: Armand R. Tanguay, Jr.)**

**A DDR&E/ARO FY 98 Multidisciplinary University Research Initiative**

**Final Progress Report**

**ARO Proposal Number: P-38565-EL-MUR**

**ARO Grant Number: DAAG55-98-1-0293**

**Research Period 1 June, 1998 – 31 May, 2004**

**Submitted To:**

**Dr. Michael Gerhold  
U. S. Army Research Office  
4300 South Miami Boulevard  
P. O. Box 12211  
Research Triangle Park, North Carolina 27709-2211**

**Submitted By:**

**Prof. Armand R. Tanguay, Jr.  
520 Seaver Science Center  
University of Southern California  
University Park, MC-0483  
Los Angeles, California 90089-0483**

## Table of Contents

Title Page	1
Table of Contents	2
Abstract	5
Executive Summary	5
Program Goals and Objectives	5
Technical Approach	6
Densely Integrated Hybrid Electronic/Photonic Multichip Module	7
Program Component Interactions	10
Interactions With Other MURI Efforts	12
Significant Accomplishments	14
Significant Accomplishments: Vision Algorithms, Models, and Architectures	14
Introduction	14
Development of an Invariant Object Recognition System: A Correspondence-Based Recognition Architecture	15
Improved Robustness of Object Recognition Using a Representation Based on Gabor-Wavelet Magnitudes	18
Modeling Object Recognition and Shape Classification: Model Robustness and Human Psychophysics	22
Recognition in the Presence of Image Variations, Particularly Pose	24
Recognition of Objects in the Presence of Variations in Distance (Scale): Learning from Human Recognition	26
Recognition of Objects in the Presence of Rotations in Depth: Learning from Human Recognition	32
Visual Representations: Analysis of Tradeoffs and Feature Learning	33
Extraction of Shape-Defining Contours	34
<b>Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures</b>	<b>2</b>
<b>Final Progress Report (1 June, 1998 – 31 May, 2004)</b>	

Extraction of Shape-Defining Contours: The PD Edge Detector	35
Extraction of Shape-Defining Contours: The Contour-Extraction Network	36
Adaptive Fusion of Cues	42
Development of a Bayesian Vision Model from the Mapping Perspective	42
Investigation of Space-Time Tradeoffs in the Adaptive Optoelectronic Eye Hardware System	47
Development of Spatio-Temporal Multiplexing Techniques	48
Enumeration of Issues for Mapping of Vision Algorithms onto the Optoelectronic Multichip Module Hardware	50
Development of a Nature/Nurture Algorithm for Visual Adaptation	50
Significant Accomplishments: Hybrid Electronic/Photonic Hardware Implementation	56
Evaluation of Dual-Input, Dual Output Silicon VLSI Neuron Unit Arrays	56
Development of a Single-Sided Flip-Chip Bonding Process	65
Development of High Refractive Index Diffractive Optical Elements (DOEs)	68
Development of High-Performance Antireflection Coatings for High Refractive Index DOEs	69
Design and Fabrication of Low Threshold Vertical Cavity Surface Emitting Laser Arrays	71
VCSEL Structure	72
VCSEL Fabrication	74
Bottom-Emitting 980 nm VCSEL Array Results	78
Photonic Multichip Module (PMCM) Integration	85
Implementation of Variable-Kernel-Size Sobel Transformations	88
PMCM Optical Power Budget	90
<b>Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures</b> <b>Final Progress Report (1 June, 1998 – 31 May, 2004)</b>	<b>3</b>



References	94
MURI Faculty Honors and Awards	97
Scientific Personnel	100
Key Faculty Investigators	100
Affiliated Faculty Investigators	101
MURI Postdoctoral Fellows, Graduate Research Assistants, Undergraduate Research Assistants, and Administrative Staff	102
Degrees Conferred	104
List of Publications	106
Journal Publications	107
Conference Presentations (Published Proceedings)	111
Conference and Technical Presentations	113
Report of Inventions	122
Technology Transfer	122
Multidisciplinary Education	125

## **Abstract**

The goal of this research program was to develop novel algorithms, architectures, and hardware for a truly smart camera, with inherent capability for semi-autonomous object recognition as well as optimal image capture. In this research, we combined striking advances in the understanding of the mechanisms of biological vision systems with similar advances in hybrid electronic/photonic packaging technology, in order to develop adaptive, artificial, biologically-inspired vision systems. A key research program objective, therefore, was to establish and address the fundamental scientific and technological issues that currently inhibit the implementation of such adaptive optoelectronic eyes. Several novel approaches to the vertical integration of multiple silicon VLSI vision chips into a hybrid electronic/photonic multichip module by means of dense 3-D photonic interconnections were pursued. In this approach, local and quasi-local connectivity between layers is accomplished by using novel diffractive optical structures that provide for both point-to-point interconnections and weighted fan-out within a local neighborhood. During this research program, significant progress was achieved in the definition of scientific and technological hurdles; the establishment of commonalities among several low-level, intermediate-level, and high-level vision models; the mapping of key functionalities onto the hybrid electronic/photonic architecture; the characterization of key components such as vertical cavity surface-emitting laser arrays (VCSELs) and diffractive optical elements (DOEs); and key steps in the integration of hybrid electronic/photonic multichip modules that are capable of implementing these vision models.

## **Executive Summary**

This section presents an overview of the completed MURI research program “Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures”, detailing the program goals and objectives as well as the technical approach, and includes schematic diagrams of the envisioned adaptive optoelectronic eye module.

## **Program Goals and Objectives**

The development of an intimately coupled sensor/processor module with architectural characteristics and capabilities similar to those found in the multilayer retina and early stages of vision in the mammalian visual system was a primary program goal, and represents the instantiation of biologically-inspired vision models, algorithms, and architectures in a compact multilayer, vertically integrated, multichip module that provides both sensing (image acquisition) and processing functions (for image recognition and feature extraction, for example). The proposed intimate coupling of the sensor and processor is a key differentiating feature, and has close biological analogs.

Preprocessing optics are also envisioned for pre-focal-plane image acquisition efficiency (to solve the traditional fill-factor problem), color

differentiation, resolution pyramid pre-image-structuring, and mapping of both foveal and peripheral regions of the sensor field-of-view [Veldkamp, 1993].

The development of a hybrid electronic/photonic multilayer chip packaging technology, integrating hybrid analog/digital VLSI chips with dense optical interconnections was also an important program goal [Tanguay and Jenkins, 1996]. This research program proved to be an excellent test vehicle for demonstration of the added value provided by the emerging photonic technology toolbox in compact image processors, and for the elucidation of remaining unresolved scientific and technological issues in the hybrid integration of optical and photonic components with mixed representation analog/digital VLSI technology. In addition, the development of this integrated hybrid packaging technology has potential for a wide range of applications beyond those that are vision-related (for example, in optical fiber communications).

The incorporation of environmental adaptivity (both short term and long term) into these adaptive optoelectronic eye modules represents the capability for optimization of image acquisition (for example, under different illumination conditions) as well as for optimization of target recognition (for example, based on the most recently observed examples of a given type of target, or taking into account the changes in image acquisition conditions (such as lighting) since the last observation). As such, this program component represents the investigation of how to best incorporate learning (based on both short term and long term memory) into a combined sensor/processor architecture that is thereby adaptive to the local environment.

## **Technical Approach**

One key aspect of the technical approach is the development, test, and evaluation of biologically-inspired vision algorithms and architectures that extract key features from existing biological paradigms, but at the same time respect the differences in capabilities, characteristics, and implementation densities between the original biological toolbox (wetware) and the emerging hybrid electronic/photonic toolbox (electronic/photonic hardware).

Therefore, a second key aspect of the technological approach is the mapping of vision algorithms onto hybrid electronic/photonic hardware, such that the performance successes demonstrated by the developing vision models on workstations with relatively long computation times can be replicated on an advanced technological substrate with far shorter computation times (greatly reduced latency and increased frame rate capability). The major thrust is to start with the biological paradigm (implemented in wetware), extract key conceptual algorithms and architectures for certain visual tasks, demonstrate the effectiveness of these models (implemented in software at the workstation level) by direct comparison with results obtained using human observers, and then re-map the key algorithmic and architectural features (such as dense fan-out/fan-in interconnections in a multilayered configuration) onto multiple planes of VLSI

chips interconnected with dense optical/photonic/electronic fan-out/fan-in interconnections.

Finally, the third key aspect of the technical approach is to design, fabricate, and test densely integrated, compact, low-power 3-D electronic/photonic multichip modules (MCMs) that incorporate the desired functionalities of adaptive optoelectronic eyes for a wide range of applications.

### **Densely Integrated Hybrid Electronic/Photonic Multichip Module**

The basic architectural concept envisioned for an adaptive optoelectronic eye, based on a densely integrated hybrid electronic/photonic multichip module, is illustrated schematically in Fig. 1. This conceptual diagram shows the integration of multiple VLSI chip layers that are optically interconnected with dense fan-out/fan-in interconnections between each of the layers.

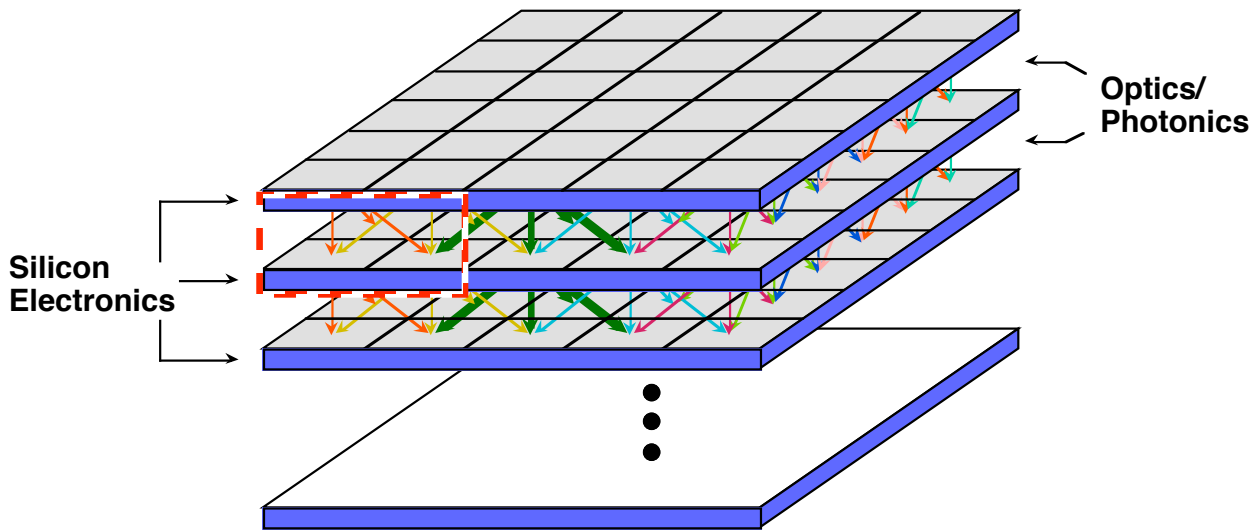
The example shown represents a pixellated structure in each layer of the sensor/processor stack, with each layer comprising a 2-D array of processing elements. Each such processing element may include, for example, analog neuron-like signal processing functions, digital functions such as sample-and-hold, local memory, and in some cases may also contain communication between neighboring pixels as well as additional control functions. Specific configurations in each layer may also include analog, digital, or hybrid analog/digital mixed representation processors; shifting circuitry for lateral scrolling functions; and the actual sensor elements (photodetectors and preamplifier circuitry) themselves.

Figure 2 schematically illustrates a cross-sectional view of one possible implementation of the dense optical fan-out/fan-in interconnections between each pair of layers in the multichip module, and hence illustrates the region identified by the highlighted rectangle in Fig. 1 (described above). In this case, the silicon VLSI chip combines both detectors (illuminated with image-bearing or previous layer information from above in a through-substrate configuration) and processing electronics within each pixel [Jenkins and Tanguay, 1992]. A multiple quantum well (MQW) modulator array is shown flip-chip bonded to the silicon pixel array on a pixel-by-pixel basis (with either three or four flip-chip bonds per pixel, in particular for the case of dual modulators representing both positive and negative output signals, comprising two signal channels and at least one if not two ground connections).

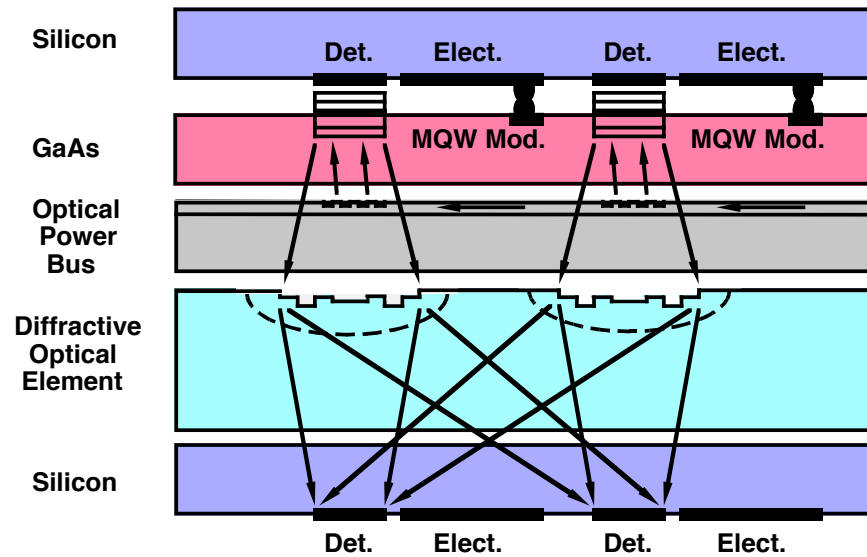
An optical power bus (an integrated optical component) delivers pixellated readout beams to each modulator element in the array by employing a combination of a rib waveguide array with vertical outcoupling gratings that are fabricated on top of each rib waveguide within the array. The center-to-center spacings of the rib waveguides and of the outcoupling gratings as well are designed to match the pitch of the individual modulator elements.

The diffractive optical element (DOE) array is designed to provide weighted fan-out interconnections from each individual modulator element in the upper

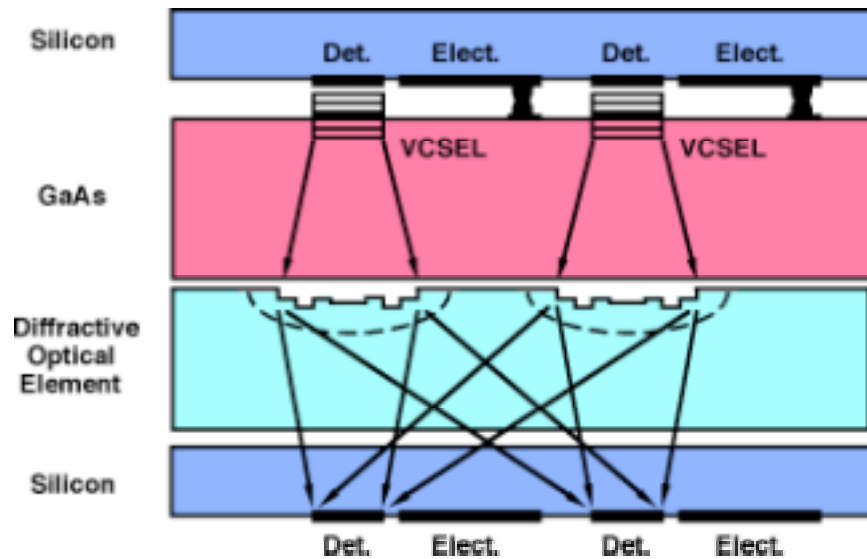
silicon VLSI chip to a number of neighboring photodetectors within pixels located on the lower silicon VLSI chip, also back-side illuminated. Appropriate focal power can be included either within the design of the DOE array itself (at the cost of space-bandwidth product), or by incorporating an in-diffused refractive microlens array, or by incorporating a separate microlens array fabricated on an additional substrate.



**Fig. 1.** Conceptual diagram of 3-D optoelectronic structure, showing silicon analog/digital VLSI chips and optical fan-out/fan-in interconnections.



**Fig. 2.** Schematic diagram of a multilayer hybrid electronic/photonic computation/ interconnection element, showing the novel optical power bus, as well as the MQW modulator and diffractive optical element arrays.



**Fig. 3.** Schematic diagram of a multilayer hybrid electronic/photonic computation/ interconnection element, showing the vertical cavity surface emitting laser and diffractive optical element arrays.

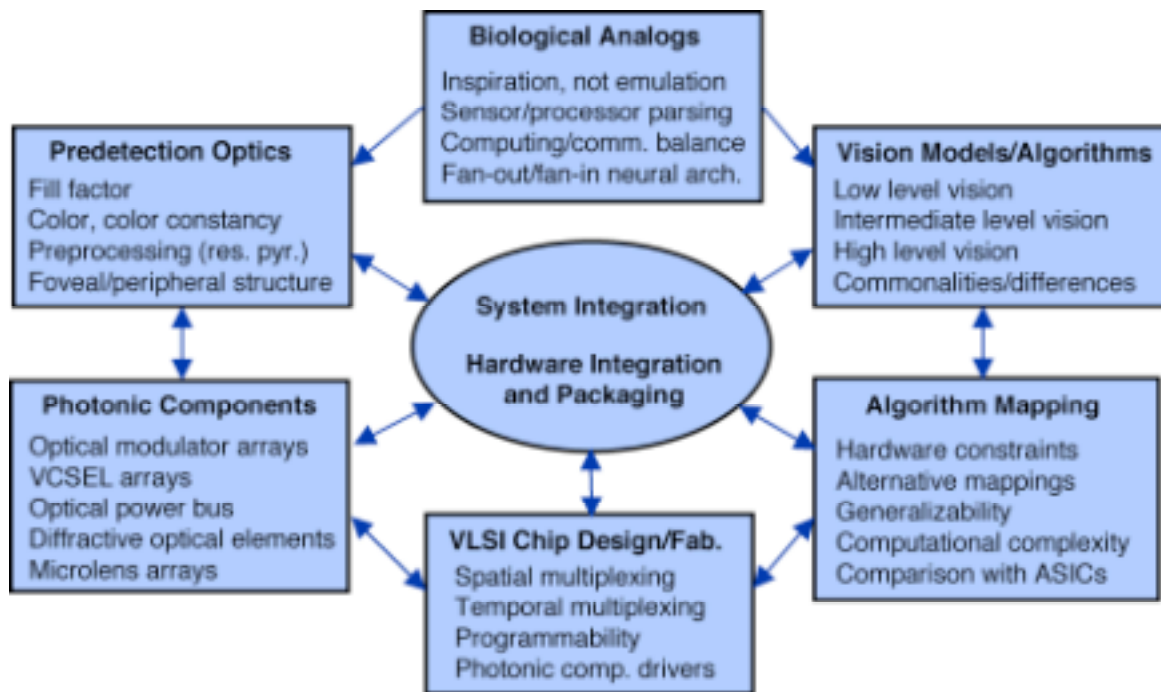
Figure 3 schematically illustrates a cross-sectional view of another possible implementation of the dense optical/photonic/electronic fan-out/fan-in interconnections between each pair of layers in the multichip module. In this case, the optical power bus and 2-D MQW modulator array have been replaced by a 2-D vertical cavity surface-emitting laser (VCSEL) array, shown flip-chip bonded to the silicon pixel array on a pixel-by-pixel basis (also with either three or four flip-chip bonds per pixel, as described above). The use of VCSEL arrays simplifies the architecture and eliminates the need for the optical power bus. However, as will be discussed further below, the use of VCSEL arrays is currently restricted to either sparse arrays or low duty cycle operation as a result of power dissipation limitations.

### **Program Component Interactions**

The various key program components and subcomponents as well as the interactions among the various elements are shown schematically in Fig. 4. All of the key program components are focused on the development of an adaptive optoelectronic eye, and as such are defined by a combination of system integration issues and hardware integration and packaging issues.

The critical importance of multidisciplinary contributions is evident in this figure, as no one faculty member's expertise spans such a wide range of scientific and technological approaches, including experience with the capabilities of the human visual system (including access to and protocols for human subjects); theoretical capabilities; system analysis and modeling skills; breadth of simulation tools and computer-aided design tools; experience with device design, characterization, and testing; and system-level (or sub-system-level) experimental facilities.

The academic disciplines represented by the various faculty members range from Neuroscience, Computational Neurobiology, and Psychology to Biomedical Engineering, Computer Science, Electrical-Engineering-Electrophysics, Electrical Engineering-Systems, Materials Science, Chemical Engineering, and Physics, as shown in the Scientific Personnel Section below. The collective academic expertise within the MURI team spans the psychology of vision, the physiology of vision, neurobiology, computational neuroscience, neural networks, the development and modeling of vision algorithms, VLSI device design and fabrication (both analog and digital, including photosensor arrays), optical device design and fabrication (including diffractive optical elements, stratified volume holographic optical elements, integrated optical devices, rib waveguide arrays, and optical power buses), photonic device design and fabrication (including both 2-D MQW modulator arrays and vertical cavity surface-emitting laser arrays), hybrid electronic/photonic packaging, and flip-chip bonding.



**Fig. 4.** Schematic diagram of the key components of the research program, depicting the inherent program multidisciplinary as well as the interactions among the various components.

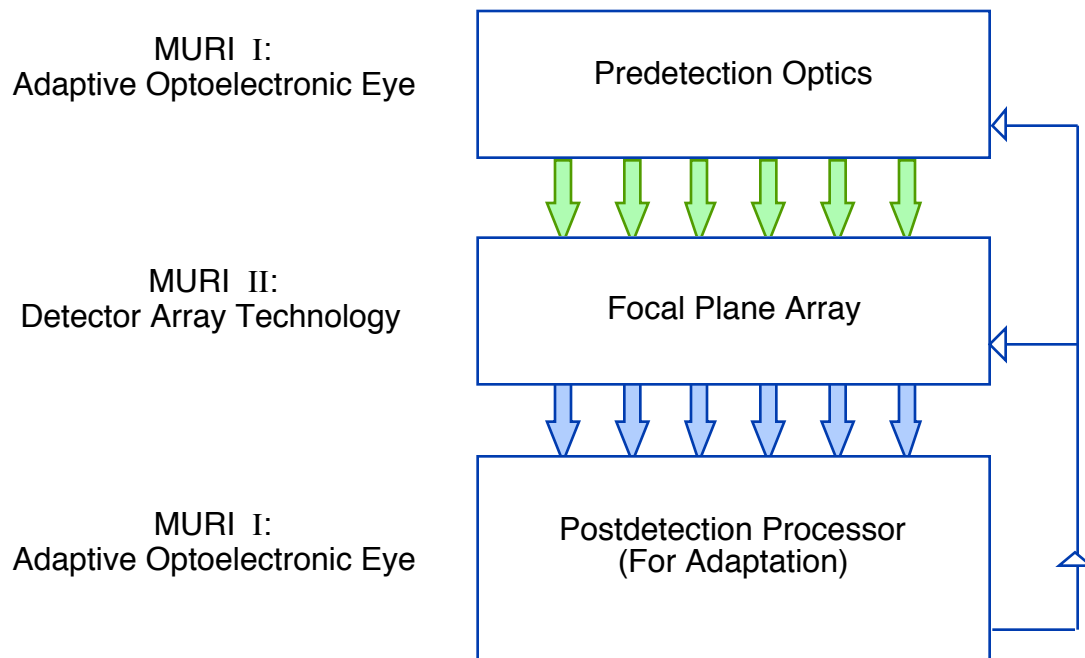


## **Interactions With Other MURI Efforts**

The envisioned adaptive optoelectronic eye sensor system is targeted for image information in the visible spectrum, within the spectral sensitivity range of the first (top-most) silicon photodetector/processor array, as shown schematically in Figs. 1, 2, and 3. However, it is of considerable interest to extend the capabilities of such adaptive optoelectronic eyes to complementary spectral regions, such as the mid-, near-, and far-IR regions. As such, the focal plane array envisioned in the current effort can conceivably be replaced by an appropriate sensor array that is optimized for one or more of these spectral regions, as shown in Fig. 5.

To this end, we developed a strong interaction with a second MURI effort that was focused on IR detector arrays based on emerging quantum dot technology ("Stress-Engineered Quantum Dots for Multispectral Infrared Detector Arrays", FY 98 MURI Program, Contract No. F49620-98-1-0474; Principal Investigator: Prof. Anupam Madhukar, University of Southern California; Program Manager: Maj. Daniel K. Johnstone, Air Force Office of Scientific Research). The goal of this related research program was to develop IR focal plane arrays with enhanced sensitivity and quantum efficiency by making use of the significant increase in absorption cross section that results from 2-D quantum confinement.

As shown in Fig. 5, predetection optics developed under MURI I (Adaptive Optoelectronic Eyes for short) and redesigned for use in the infrared focus and disperse the incoming image illumination onto a focal plane array that incorporates stress-engineered quantum dot detectors that are both pixellated and multi-wavelength, as developed under MURI II (Detector Array Technology for short). Outputs from the quantum dot detectors are fed in parallel to a postdetection processor based on the technology described herein (incorporating a hybrid electronic/photonics multichip module that implements advanced adaptive detection and vision-related functions), again as developed under the current MURI I program. Coupling between the IR focal plane array and the postdetection processor is envisioned to comprise the same type of dense flip-chip bonding that is described herein as key to the MURI I effort.



**Fig. 5.** Schematic diagram of an augmented adaptive focal plane array system. A set of predetection optics modifies the incoming image before it is detected, potentially increasing detection efficiency and color discrimination. Adaptive, localized gain control is provided to the focal plane array, in turn providing improved utilization of detector dynamic range and enhancing detection of objects in the presence of nonuniform illumination. The postdetection processor provides on-board computation of these adaptation signals, as well as region-of-interest localization for pointing and zoom control (if appropriate). The primary research focuses of the two interacting MURI efforts described in the text are also shown herein.

## **Significant Accomplishments**

This MURI grant was awarded under the FY-98 Multidisciplinary Research Program of the University Research Initiative, with a start date of 1 June, 1998. In the sections that follow, a summary of significant accomplishments during the research program period (1 June, 1998 through 31 May, 2004) is provided.

Accomplishments in the general area of vision algorithms, models, and architectures, on the one hand, and in the general area of hybrid electronic/photonic hardware implementations, on the other, are described separately below for clarity. Extensive interactions between these two principal components of the efforts proved crucial to the success of the research program.

### **Significant Accomplishments: Vision Algorithms, Models, and Architectures**

#### **Introduction**

The main difficulty in the development of vision algorithms, models, and architectures stems from the tremendous variation of natural scenes. No two scenes are ever alike in any superficial sense. Even a concrete, individual object never looks the same in two images. This variation defies rigidly constructed vision mechanisms, although this is the traditional approach of computer vision. In order to make progress, we can learn some valuable lessons from biological vision systems. One of these lessons is the importance of adaptivity on all time scales. Besides the slower time scale of learning from example, it has become more and more evident that adaptation is also necessary (and taking place in animal visual systems) during the process of scene interpretation, for instance to retune individual visual cues in the light of information from other cues.

It has been the classical style of computer vision to “construct” vision mechanisms on the basis of physical and mathematical insight. It is becoming increasingly clear that the battle cannot be won in this manner. Biology teaches us that “learning from examples” is a much more successful strategy. Once we have developed techniques to apply this style on all levels of the visual hierarchy, from low-level features via feature arrangements, to object parts, whole objects, and to entire scenes, the rest of the work can be done automatically by adaptive visual systems of which only the general architecture has been constructed. To make progress towards this architecture, both on the hardware and software level, was the central goal of our project.

Thus, our vision algorithm, model, and architecture effort addressed, in part, two key problems facing the implementation of a useful and generalizable eye/vision module: first, the recognition of scenes and objects in a manner that is robust with respect to typical image-to-image and object-to-object variations, which include rotations (laterally and in depth), scaling and translation, occlusion, scene illumination, and distortion of objects; and second, the implementation of these algorithms in parallel hardware that provides

sufficiently fast computation to yield results in at least real time [Tanguay, *et al.*, 2000].

Toward these ends, we designed the research effort to cope with several levels of the visual hierarchy. On the lowest level, an effort led by A. R. Tanguay, Jr. and B. K. Jenkins focused on developing the technology to implement feature extraction (*e.g.*, of Gabor-type wavelets) in an efficient parallel, partially optical, partially or wholly analog technology, based on a paradigm first described in [Tanguay and Jenkins, 1996]. On a higher level, an effort led by B. Mel examined families of mid-level feature types and combinations of features, among them edges and edge combinations, *e.g.*, [Mel, 1997]; this work includes the consideration of techniques for learning efficient and useful sets of feature types from examples. In addition, an effort led by C. von der Malsburg made strides in the development of representations of objects that are particularly robust in the presence of changes in such parameters as pose and distortion. On the level of object components (geons), an effort led by I. Biederman exploited the rich treasure of biological information, gleaned by his group and interpreted using their psychophysical methods, in order to better model the object recognition process of the human visual system, thus providing us with pertinent biological inspiration.

On still higher levels of whole objects and scenes, it is necessary for artificial and natural visual systems to absorb enormous quantities of sample material (children do this for more than a decade before their visual systems reach full competence), and it is necessary to both condense that sample material as much as possible and ready it for generalization, interpolation, and extrapolation. In addition it will be necessary to develop means to merge that domain knowledge in an efficient way with visual inputs in order to decipher them. Efforts led by C. von der Malsburg made progress on that front by constructing statistical object models that are parameterized by such variances as pose or deformation, and by developing efficient matching mechanisms to compare such models to image components.

Finally, we also describe below progress in our effort to achieve techniques for mapping useful vision algorithms and models onto the photonic multichip module hardware, in order to implement them in a fast, highly parallel, adaptive module.

### **Development of an Invariant Object Recognition System: A Correspondence-Based Recognition Architecture**

One key element of the vision of this collaborative project was to recreate the fantastic functionality of biological vision systems as a new technology. This is a tremendous challenge on two levels. On the one hand, we had to come up with clear concepts of a system architecture, and on the other hand an appropriate implementation technology had to be developed. These two issues are not independent of each other at all, and required close collaboration and intensive discussion. The biological model is not perfectly understood itself, and even if it was, the constraints under which electronic and photonic technology

operate are fundamentally different from those reigning in biology, so that creative approaches are required.

One of our key strategies in the context of this research project was to develop an invariant object recognition system that is competitive in the technological domain and at the same time can be taken seriously as a model of the biological vision system. It is our conviction that, if done right, this system will act as a paradigm for vision in general. "Doing it right" means addressing a number of critical aspects:

#### **Attention Control**

The system has to select subregions in the image or scene that are of interest in the context of current system goals. This can be driven bottom-up (by features, such as local movement, that signal significance) or top-down, in a search mode, to select from among the candidates offered bottom-up. Top-down attention control already presupposes a system for object representation and for mapping between an invariant domain and the image domain, as described further below.

#### **Figure-Ground Separation**

The object at the focus of attention needs to be clearly separated from the background and from potential occluders. This is especially important if the contour of the object is to be extracted with some precision.

#### **Invariance**

Depending on viewing conditions, the image of an object in an eye or camera is subject to variance in terms of position, size, orientation, pose (rotation of the object in depth), deformation, surface marking, illumination, partial occlusion, background, and noise.

#### **Feature Definition**

From the images as delivered by a camera or the eye, features are to be extracted that are sensitive to the differences that are important during recognition, and insensitive to differences that are irrelevant.

#### **Object Representation**

For many applications, it is important to have explicit representations of the structure of objects (admitting that the mere classification of objects may be possible without such, based on appropriate features).

#### **Subsystem Integration**

For object recognition, subsystems for shape, surface markings, contour form or object parts (to name a few examples) need to be combined. In figure-ground separation, subsystems have to deal with motion, texture, color, form, and contour shape, among other features.

#### **Single-Example Learning**

A system must be able to learn from single examples of a new object.

## Speed

The system needs to be highly parallel. The biological system recognizes objects within the time of a few neural transmission delays.

This project, combined with support from other sources, permitted our research effort to make decisive progress on essentially all of the above aspects. The basis of the architecture that has emerged is a system for correspondence-based recognition. The system is composed of two domains, an image domain  $I$  and a model domain  $M$ . The image domain has the form of a two-dimensional sheet of nodes (in the biological version, the nodes are local sets of neurons of different specificity). Nodes express information on local features. In the conceptually simplest version of the system, the model domain consists of a collection of sheets of nodes, each very similar to the image domain, one such sheet per recognizable object. The task of the system is to find a correspondence map, that is, a set of links between points in  $I$  and models in  $M$ . In an acceptable correspondence, there is high similarity between corresponding points in  $I$  and  $M$ , and neighboring nodes in  $I$  map to neighboring nodes in  $M$ .

One of the key thrusts of this part of the overall project was the optimization of the object recognition performance of the system. This led to publications [Okada, *et al.*, 1998; Triesch and von der Malsburg, 2001; Triesch and von der Malsburg, 2002; Serrano, *et al.*, 2003; Huesken, *et al.*, 2004; and Eckes, *et al.*, 2006]. In [Okada, *et al.*, 1998], it was shown in competitive tests with other leading groups that our correspondence-based approach combined with Gabor wavelet features is highly competitive. In [Triesch and von der Malsburg, 2001; Triesch and von der Malsburg, 2002], we tested the system on another type of objects, human hands, whose posture was to be recognized. In [Serrano, *et al.*, 2003], we explored possibilities for compressing object representation data under the constraint of undiminished recognition rate. In [Huesken, *et al.*, 2004], we devoted some effort to improved face recognition rates in the presence of head pose differences. In [Eckes, *et al.*, 2006], we recently showed that the system is able to analyse cluttered scenes and direct its attention sequentially to different objects.

A second thrust of this project was directed at better understanding of Gabor wavelet features. These are directly inspired by the biological model. One of their virtues is their robustness against small object deformations and small image shifts. The latter necessitates, however, abolishing the Gabor phase, by taking the magnitude of the signal (square root of the sum of squares of the sine and cosine responses). We were able to show [von der Malsburg, *et al.*, 1998; Shams and von der Malsburg, 2002; Wundrich, *et al.*, 2004] that the only ambiguity incurred by abolishing phase information is that between the original and the photographic negative. This research direction is described in further detail in a subsequent section. In [Kalocsai, *et al.*, 2000], we studied the relative contributions of different Gabor components to the recognition success, and found that some components are more important than others by factors of several thousand.

Central to the success of the comprehensive project was coming up with a

mechanism of rapid correspondence map formation. In technical implementations, as now used for face recognition in several leading companies, a number of trial maps are tried out in sequence, using first large, then smaller and smaller steps in parameter space (e.g., relative position), finally permitting the map to deform slightly, while optimizing the overall similarity between correspondence points in the image and the model. As a biological model, this is unacceptable; in the technical domain, this is too slow and ill-adapted to parallel implementation. An early version [Wiskott, *et al.*, 2002] may apply in the biological case early in ontogenesis (and was and is discussed by us as a possible technical implementation); it is not realistic, however, to explain object recognition in the adult animal, being too slow by orders of magnitude. The basic issues involved are discussed in [von der Malsburg, 2002a; von der Malsburg, 2002b].

In a series of papers [von der Malsburg and Zhu, 2000; Zhu and von der Malsburg, 2001; Zhu and von der Malsburg, 2002a; Zhu and von der Malsburg, 2002b; Zhu and von der Malsburg, 2003; Zhu and von der Malsburg, 2004], we were able to show, however, that if there was a mechanism for direct interaction between links, very rapid map establishment and recognition is possible (within a few iterations). That early implementation of the dynamic link mechanism still suffered from some technical difficulties (feature similarities had to be computed off-line). In [Luecke, *et al.*, 2002] we overcame those difficulties with the help of a macrocolumnar realization of nodes in the image and model domains. We have recently shown that object recognition can be realized in a fully parallel system [Luecke and von der Malsburg, ICANN06, submitted]. Under separate funding, we are now collaborating with R. Douglas, Zurich, on a VLSI implementation of that system, thus coming very close to one of the original goals of the project reported here.

A final thrust partially supported under this project was directed at learning. In [Luecke and von der Malsburg, 2004] we developed a novel system for feature learning, with which it has been possible since then to show learning of Gabor wavelets from natural inputs. In [Prodoehl, *et al.*, 2003] we showed that feature relations can be realistically learned from natural image sequences. These relations are important for the extraction of contours from images, an ability highly developed by our colleague B. Mel within this project. In [Shams and von der Malsburg, 1999; Shams and von der Malsburg, 2002] we were able to show learning of shape primitives (geons), as described in psychophysical experiments by our colleague I. Biederman. In [Zhu and von der Malsburg, 2002; Zhu and von der Malsburg, 2006] we were able to demonstrate generation by learning of the link-to-link associations that are necessary for rapid correspondence finding. Finally, in [Loos and von der Malsburg, 2002] we demonstrated one-shot learning of object models from natural visual scenes.

### **Improved Robustness of Object Recognition Using a Representation Based on Gabor-Wavelet Magnitudes**

One of the most fundamental and difficult issues of image understanding and visual object recognition is the attainment of invariance or robustness to

variations in the appearance of objects and scenes from image to image. The geometrical variations of translation, scaling, and rotation within the image plane can be dealt with in a systematic and comprehensive way. Other changes, however, cannot be modeled practically. Among these sources of variation are changed illumination, object distortion, and (at least in the absence of precise 3-D shape information) changes in perspective.

Described herein is work on the use of Gabor-wavelet magnitudes (without phase information) to represent an object. Evidence is given that the use of these Gabor magnitudes provides robustness with respect to many parameters, including slight variances in pixel positions of features within an object or scene, and to contrast reversals of parts of the object. A vision model based on Gabor-wavelet representation and elastic graph matching, Christoph von der Malsburg's dynamic link architecture model [von der Malsburg, 1981], has been used very successfully in the past to model faces and to recognize people by pictures of their faces. This new work using Gabor magnitudes only, without the phases, allows for increased robustness of the model, and has the potential to significantly reduce the complexity of instantiating vision models of this kind in the emerging hybrid electronic/photonics hardware.

Von der Malsburg's dynamic link architecture vision model uses graphs labeled with Gabor-based wavelets to represent objects and recognize them by elastic graph matching in a way that is invariant to the geometrical transformations listed above, and robust with respect to deformations [Lades, *et al.*, 1993; Wiskott, *et al.*, 1998]. Gabor wavelets are two-dimensional sampling functions (receptive field profiles or convolution kernels in biological or mathematical parlance, respectively) in the form of sinusoidal waves multiplied by suitably scaled Gaussian envelope functions. Gabor wavelets can differ both in orientation and in scale (or spatial frequencies); we typically sample the image on five scales and eight orientations. Each individual kernel type comes in two varieties, sinusoidal or cosinusoidal, depending on whether the wave is centered on the Gaussian with its zero or its maximum, respectively. From this pair of numbers (one from the sine kernel and one from the cosine kernel), which can be extracted from a given kernel type and on each image point, an equivalent pair of numbers can be computed: the magnitude (as the square root of the sum of the squares of the sine and cosine component) and the phase (as the arctangent of the ratio of sine and cosine components).

Gabor phases contain important information as to the location of image features (especially of edges). However, it turns out that the inclusion of phases is cumbersome for elastic graph matching, leading to many distracting local optima. When ignoring Gabor phases and working with Gabor amplitudes alone, on the other hand, elastic graph matching proved in our hands to be very successful for object identification purposes [Wiskott, *et al.*, 1998]. However, prior to the work reported here it was unclear what ambiguities arise when phases are ignored, and correspondingly what false matches to wrong objects or patterns are to be expected.

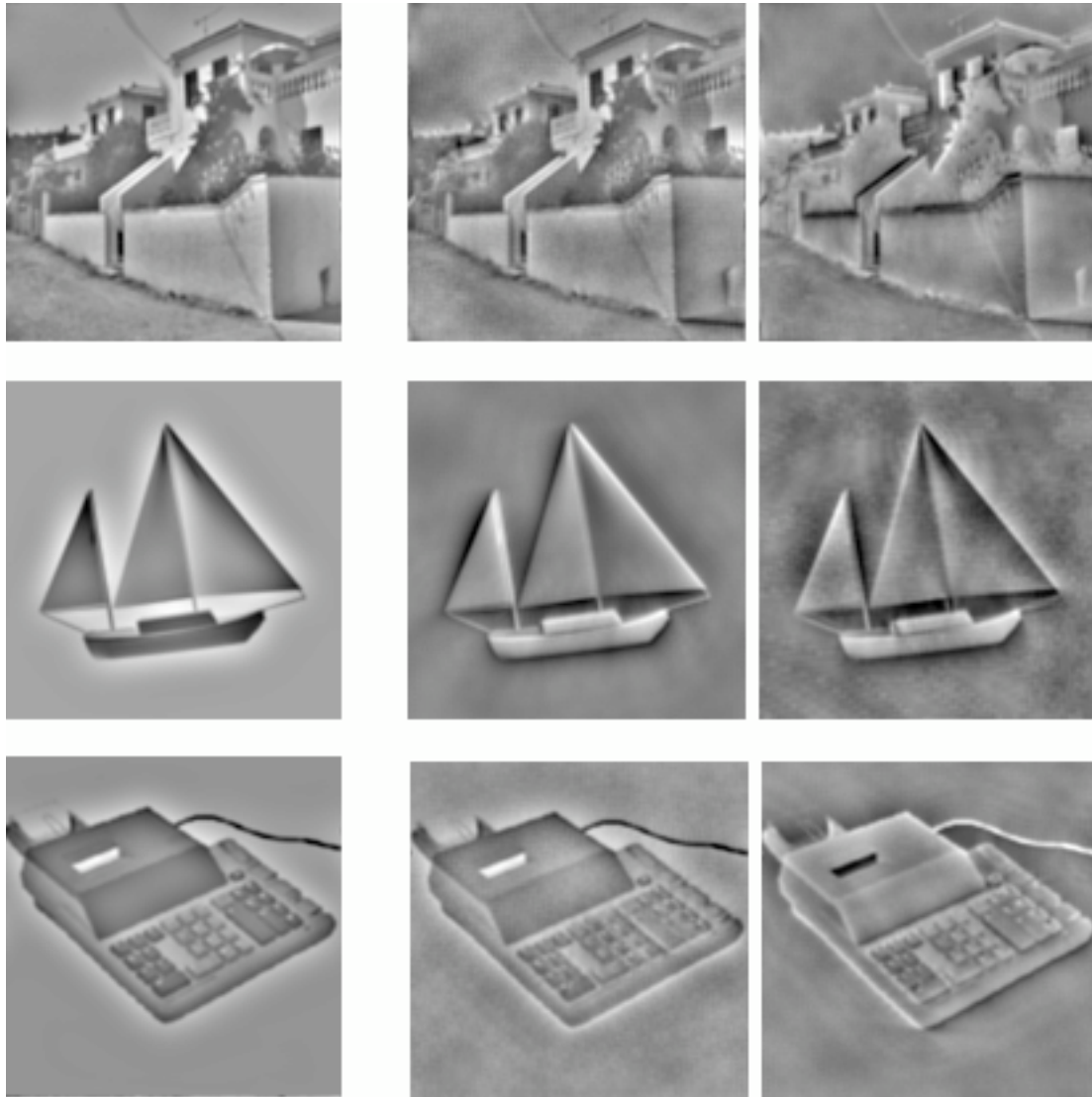


We are now able to characterize the ambiguities of image representation by Gabor magnitudes. By applying a theorem [Hayes, 1982] about ambiguities of Fourier magnitudes, we were able to prove that when the recorded Gabor magnitudes are precise (in terms of amplitude and recorded position), the only image ambiguity left is that of global contrast inversion (which would let a positive image be confused with its photographic negative). However, if the limited precision of actual image data is taken into account, a much wider range of image ambiguities is opened, and this range proves to be a boon rather than a disadvantage. A theoretical argument shows that these ambiguities correspond to small local image distortions (especially, small shifts in the position of edges) and to contrast inversions of local regions of the image.

In order to corroborate this argument we developed an algorithm to reconstruct images from Gabor magnitude information only. The algorithm starts from an arbitrary seed image, which is iteratively modified to reduce the deviation between the Gabor magnitudes of a target image and the reconstructed image. We show that this deviation can be made arbitrarily small. In accordance with the theoretical result, all accurate reconstructed images are either a detailed reconstruction of the original pixel values, or a reconstruction of the negative of the original. However, if the reconstruction procedure is stopped when the median error in Gabor magnitudes is a few percent, an interesting variety of reconstructed images is obtained. (Examples are shown in Fig. 6.) In these images, local contrast can be reversed in some regions and not in others, edges change their polarity or their profile (even such that edges are sometimes changed into lines or vice versa), and close scrutiny shows that the exact position of edges is sometimes shifted. Thus, both theoretical argument and experimental work show that there is a variety of images that agree in Gabor magnitudes up to a certain precision but differ considerably in pixel values.

It is a remarkable fact of high biological significance that the objects or scenes can always be easily recognized by our eye from the reconstructed images (indeed it is very difficult to remember differences between reconstructed images). In an ongoing collaboration with Prof. Biederman in the context of this MURI research program, we were able in psychophysical experiments to show that object identification is not affected by local contrast reversals at all. Preliminary results are reported in [Subramaniam and Biederman, 1997]. Our visual system contains as a very important component a class of cells ("complex cells") whose response characteristics are those of Gabor magnitudes, showing that evolution has found it advantageous to use them as well [Shams and von der Malsburg, 2002].

The functional significance of the stated properties of Gabor magnitudes, especially relevant to a number of military and civilian applications, is that the range of actual image variations that a recognition system has to deal with are of the kind to which Gabor magnitudes are insensitive. Among these are small distortions of objects (due to perspective changes or intrinsic distortion), changes in lighting (which can reverse the contrast of edges) or, in the case of infrared images, changes in the temperature profile of an object.



**Fig. 6.** Reconstruction of images from Gabor magnitudes. Left column: target images. These images are filtered such as to contain all and only such information as captured by our set of Gabor kernels. Middle and right columns: reconstruction obtained using different seed images. Note the local contrast inversions and changes in the profiles of rendered edges. Note also that the objects or scenes can be effortlessly recognized by our eye.

## **Modeling Object Recognition and Shape Classification: Model Robustness and Human Psychophysics**

A more fundamental way to test model robustness (*i.e.*, to assess how robust the model is with respect to the appropriate image variations, and to the appropriate degree) is to compare it to the performance of human beings on similar tasks. We have also performed experiments on von der Malsburg's vision model that compare its performance with that of trained observers, with results that show striking similarity in performance, as described in more detail below.

During the research program, we addressed the question of whether a neurally inspired Gabor-kernel based model, specifically von der Malsburg's dynamic link architecture, provides a promising basis for modeling shape recognition, by performing tests that compare the robustness of the model with that of human subjects. The general methodology we used in this set of experiments was to measure recognition performance under difficult viewing conditions (*e.g.*, brief exposures, noise, low contrast, small size) designed to lead to variations in the speed and accuracy of performance.

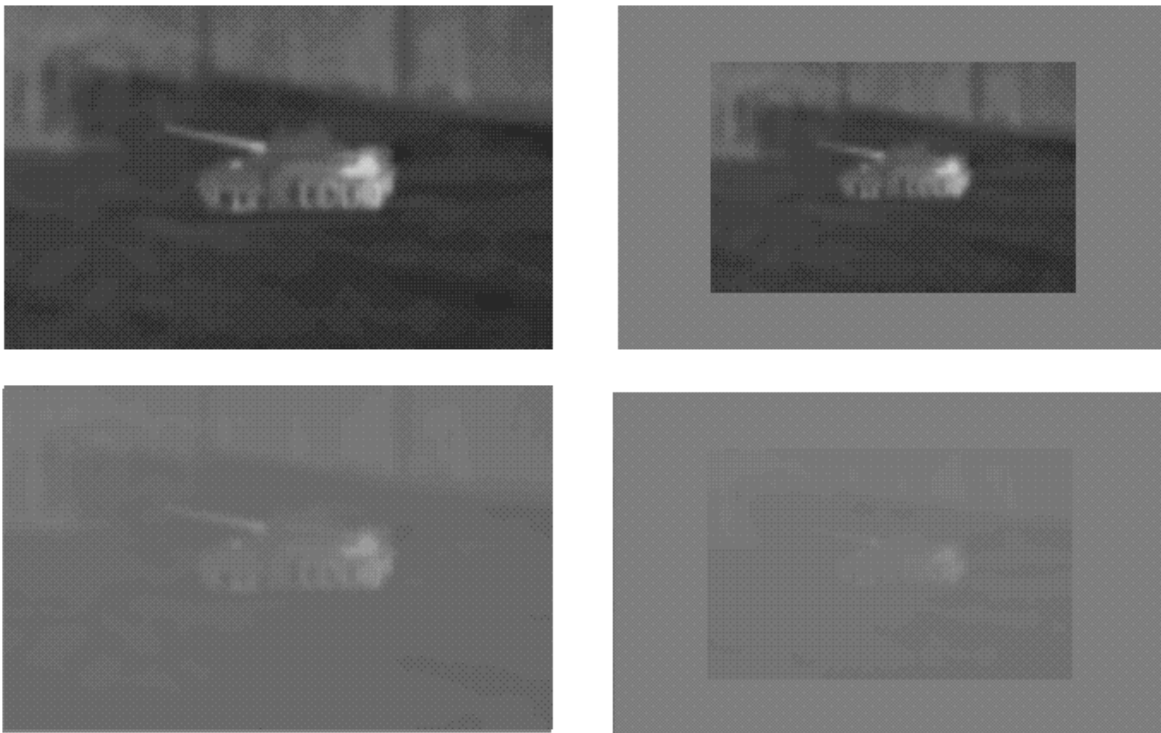
The first phase of this set of experiments was performed by Prof. Irving Biederman in conjunction with Prof. Christoph von der Malsburg, and uses FLIR images of military vehicles in realistic scenarios under variations of size (distance) and noise (*e.g.*, Fig. 7). The test assesses how often vehicle A is confused with vehicle B, resulting in a "confusion matrix". Recognition was performed by trained human observers and by von der Malsburg's vision model. Remarkable similarity (correlation of 0.89) was observed in the types of errors made by humans and in those made by von der Malsburg's model.

In further experiments, subjects performed a match-to-sample task in which they judged which of two comparison stimuli, presented briefly, was identical to the sample. The stimuli were either faces (Fig. 8) or unfamiliar, smooth, asymmetric, complex 3D blobs (Fig. 9) produced by varying the orientations of the second and third harmonics of a sphere and then adding these orientations to the sphere and fourth harmonic. The (dis)similarity between the matching and distracting blobs was assessed by four measures: (a) subjective pair-wise ratings made by human subjects, (b) Euclidean distances in a 2D stimulus space defined by the differences in the angles of rotation of the orientation-varying harmonics, (c) mean pixel luminance energy differences between pairs of images, and (d) von der Malsburg's Gabor-jet model (Lades, *et al.*, 1993), designed to model aspects of V1 simple-cell filtering.

The last measure is based on a wavelet-like filtering of the image by a lattice of Gabor jets, each composed of kernels over multiple scales and orientations. Similarity in the model is a function of the correlation of the activation values between corresponding kernels in corresponding jets. When matching faces (Fig. 8), the error rates correlated almost perfectly with the Gabor-jet similarity of the distractor:  $r = 0.943$ . For objects (such as those illustrated in Fig. 9), all four measures correlated positively with error rates on the match-to-sample trials:

Euclidean distance =  $-0.804$ , judged similarity =  $-0.846$ , pixel energy =  $-0.891$ , and Gabor jet =  $-0.910$ . In the absence of salient nonaccidental differences (e.g., differences in parts or whether contours are straight *vs.* curved), Gabor jets, a model based on V1 computations, does remarkably well in scaling the psychophysical similarity of faces as well as complex, novel shapes. To our knowledge this is the first time that psychophysical similarity has been successfully predicted by a theoretical model.

This humanlike behavior of an autonomous model indicates the potential of the model in other difficult areas of visual recognition. Furthermore, such comparisons against comparable human visual capabilities can provide valuable tests of the evolving vision models, as well as of their instantiations in hardware.



**Fig. 7.** Sample infrared images showing variations in size and contrast from an experiment on the identification of 15 military vehicles from infrared imagery.

An intriguing prospect that potentially accrues to such direct comparisons between human observer performance and vision model performance is the elucidation and specification of the confusion or illusion spaces that characterize both. This suggests the potential design of complementary vision models and corresponding hardware instantiations that can break the degeneracy, so to speak, and exhibit recognition capabilities that complement rather than mimic those of human observers.

### **Recognition in the Presence of Image Variations, Particularly Pose**

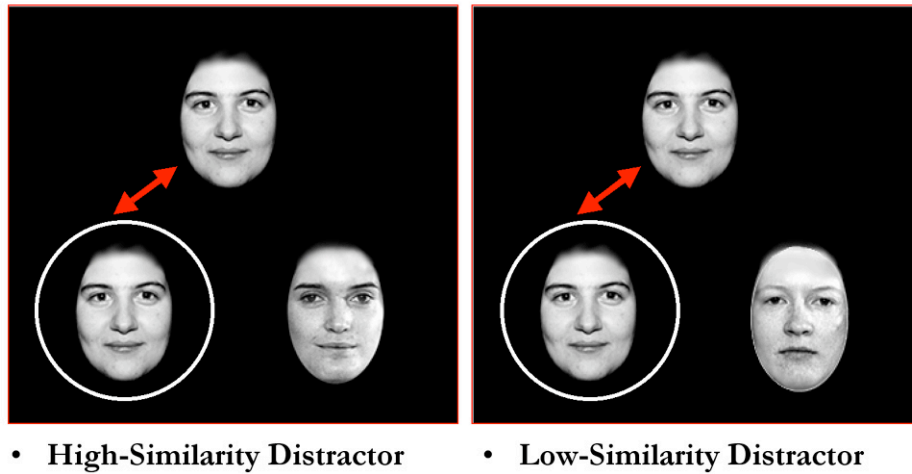
The image of a three-dimensional object varies according to position, orientation, and scale, as well as according to pose, illumination, deformation, and other sources of variation. In this section we will discuss our results on techniques for handling variations in pose that have application to automated visual recognition systems. In the subsequent two sections, we will describe several pertinent lessons we have learned from studying the recognition process in human vision systems. These lessons relate to: first, perception at various orientations in depth, which has application to changes in pose; and second, perception at various distances, which has application to changes in scale.

Our approach for many of these issues is a technique based on an object description in terms of labeled graphs, with nodes referring to points on the face of the object in the image, and with labels in the form of sets of Gabor-based wavelet components centered on the node positions. Object recognition is then implemented by a process of graph matching. In this model, the three issues of object position, in-plane orientation, and scale are dealt with explicitly and simply with the help of transformations performed on stored model graphs during the matching process, the transformation acting on node positions and attached Gabor wavelet sets.

Pose variation, on the other hand, must be handled with the help of many views from different vantage points. A simple and straightforward strategy would be to cover the viewing sphere (the domain of all pose angles) with stored aspects and search them all during the recognition process. As a background study, we have made efforts to cover the viewing sphere with as few views as possible while still being able to reconstruct all intervening views with a given accuracy as a weighted mean of neighboring stored views [Peters, *et al.*, 1999a, 1999b; Peters and von der Malsburg, 2001a, 2001b].

Working with individual stored aspects of an object is a method of limited applicability, especially when variations along more than just pose angle are at issue. We are therefore pursuing the goal of capturing variation along several dimensions in terms of a coherent object model  $X(P)$ , where  $X$  refers to the object model (the vector of data describing the labeled graph) and  $P$  is a set of parameters (*e.g.*, position, orientation, scale, pose, illumination, and deformation). If a collection of images is given that covers a small volume in  $P$ -space, principal component analysis (PCA) or independent component analysis (ICA) can be used to find a local linear basis for  $X(P)$  within the restricted

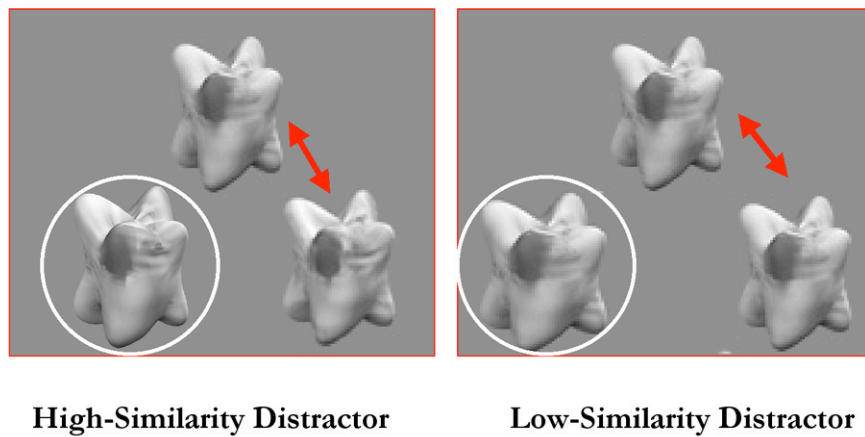
**Correlation of Gabor-jet similarity values with error rates in a match-to-sample task with faces**



$$r = .943, p < .001$$

**Fig. 8.** Illustration of match-to-sample trials on faces with distractors that are either of high similarity (left panel) or low similarity (right panel) to the matching stimulus. The greater the similarity, the greater the chance of an error.

**Correlation of Gabor-jet similarity values with error rates in matching blobs**



$$r = .910, p < .002$$

**Fig. 9.** Illustration of match-to-sample trials on blobs with distractors that are either of high similarity (left panel) or low similarity (right panel) to the matching stimulus.

volume in  $P$ -space. As  $X(P)$  will in general define a curved manifold, several partially overlapping linear descriptions have to be combined (*e.g.*, by weighted averaging) to represent  $X(P)$  over larger volumes in parameter space. We have worked with images of human heads in a range of pose angles, complete with pose angle information (provided by a magnetic tracking system). Our system allows us to recognize pose angles from a given image of a person or to reconstruct images of the person for given pose angles [Okada, *et al.*, 2000].

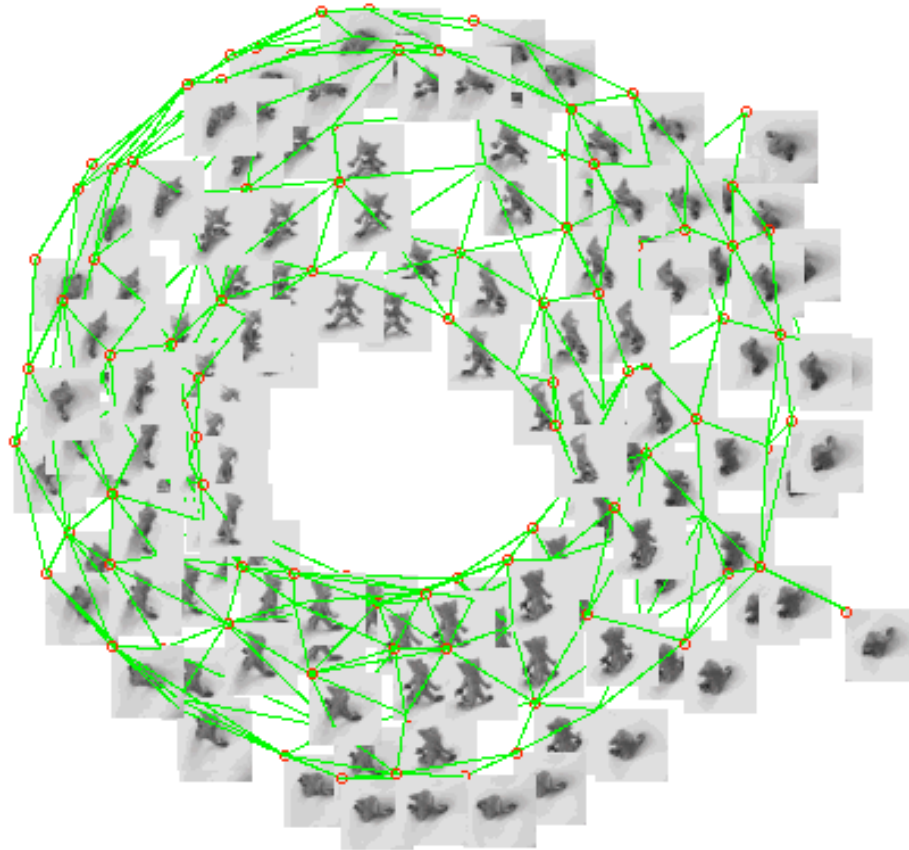
Ground truth in terms of the underlying parameter values (pose angles, for instance) is rarely given. In a further study [Wieghardt and von der Malsburg, 2000] we have taken a large set of images of a given object from different viewing angles, without those angles being given to the system. We first collected images into “viewing bubbles”: sets of images with high similarity to a central image. For each viewing bubble we erected a linear data model. We then erected a coherent three-dimensional  $P$ -space by arranging the viewing bubbles with the help of multi-dimensional scaling on the basis of pair-wise distances between the centers of partially overlapping bubbles (we derive these distances from the local coordinate frames). Finally, we aligned the linear model frames in neighboring viewing bubbles relative to each other. The resulting manifold is a coherent description of the object from all viewing angles, as shown in Fig. 10.

### **Recognition of Objects in the Presence of Variations in Distance (Scale): Learning from Human Recognition**

Two fundamental problems in the design of a visual system that can achieve shape recognition comprise how to handle variations of an image of an object when it is viewed (1) at various distances and (2) at various orientations in depth. We consider variations in distance in this section and variations in depth orientations in the subsequent section. In the case of scale, humans and animals can identify objects appearing at a variety of sizes in their visual field without much apparent cost. This problem has assumed significant import in our efforts to adapt the Gabor jet recognition system developed by von der Malsburg to human object recognition. The system assumes a spatially distributed array of Gabor filters at multiple scales and orientations that would roughly correspond to primate cortical hypercolumns in V1.

This apparent invariance over size changes poses a challenge to computational theories of visual recognition, because the early cortical representation of object features appearing at different sizes will vary greatly. For example, a slightly rounded L-shaped vertex when the object is shown at a small size will activate feature detectors sensitive to sharp curves at a given scale. The image of the same object at a larger size might not activate curve detectors at that scale at all. (L-vertices are of particular importance in segmenting the objects in a scene, as they provide a strong constraint signaling the end of a surface.)





**Fig. 10.** Representation of an object from all angles in the upper viewing hemisphere [Wieghardt and von der Malsburg, 2000]. The representation was constructed from an unlabeled set of 2500 images taken at 3.6 degree intervals. Circles correspond to center images of viewing bubbles. For each of them the corresponding view of the object is printed to the lower right. Overlapping viewing bubbles are connected by a line. The upper pole of the viewing sphere is not a point but a circle (the outer circle in the perspective figure), because during image comparison in-plane rotations were not permitted.

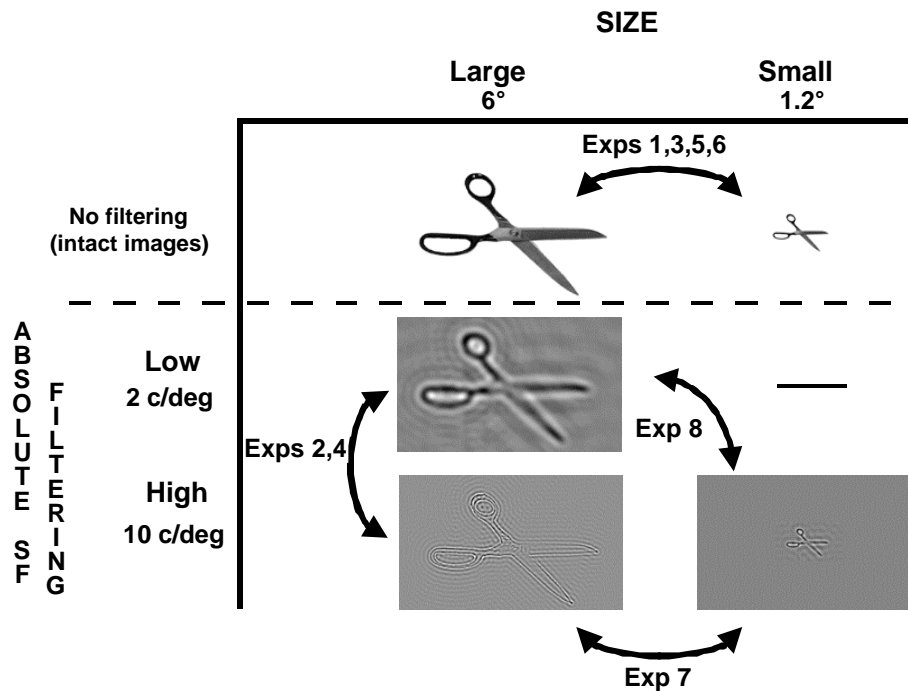


Because of an advantage in identification of low pass over high pass filtered patterns as well as large patterns over small, a number of theorists have assumed that size-independent recognition is achieved by spatial frequency (SF) based coarse-to-fine tuning. A specific proposal for implementing this idea was the assumption of a “shifter circuit”. The shifter circuit selects the most salient information on the lowest scale, and adjusts the size and sampling of the input window to higher processing centers based on this scale to achieve a size-normalized representation of a given object in the scene. Such a tuning mechanism can accommodate voluntary attentional as well as involuntary mechanisms for size and position invariant recognition.

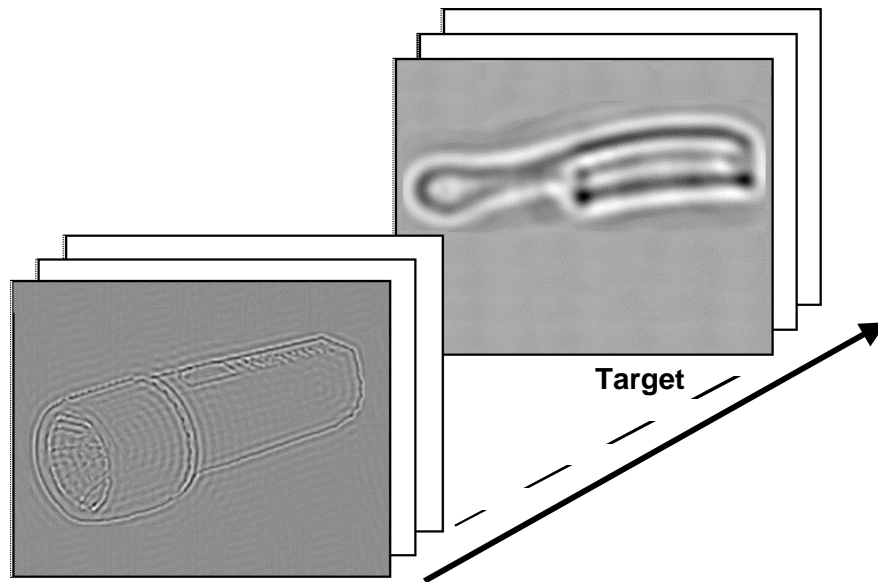
During the research program, Prof. Biederman [Fiser, *et al.*, 2001] and his co-investigators investigated this assumption of spatial frequency-based size tuning as a model of human object recognition. Specifically, they addressed the question as to whether efficient response to images of different size requires information to be represented at different scales. Would the pattern of results obtained in tasks that required processing images of various sizes be different if the spatial frequency content of the images was held restricted to different limited bands?

The experiments of Fiser, *et al.* employed the Rapid Serial Visual Presentation (RSVP) protocol, in which a time sequence of images is presented, each very briefly (*e.g.*, 72 msec). Subjects of the experiments were given a verbally specified target, *e.g.*, “chair”, and were asked to detect whether or not the target was present in the image sequence. Figure 11 shows examples of the various conditions of their eight experiments. A sample experiment is shown in Fig. 12.

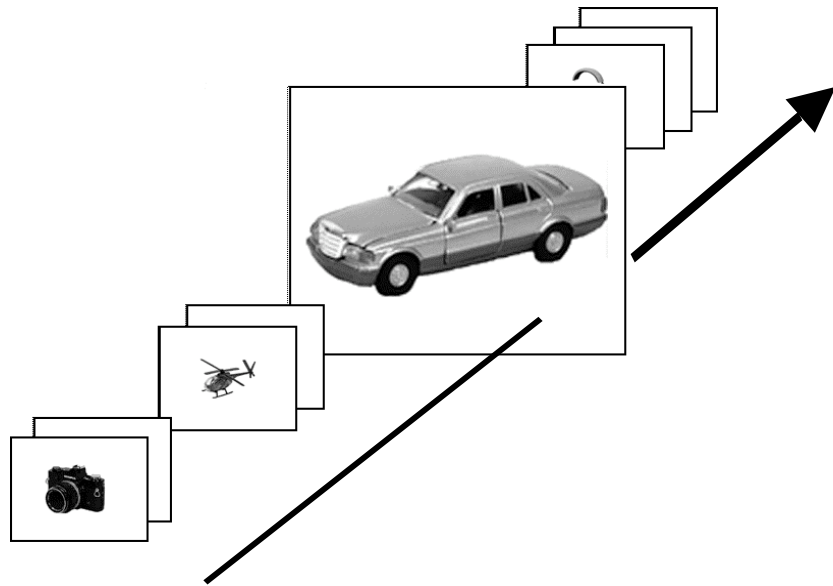
We have discovered that the advantage of large sizes or low spatial frequencies apparent in the viewing of individually presented images is lost when human subjects attempt to identify such a verbally-prespecified target object somewhere in the middle of a sequence of 40 images of objects (each shown for only 72 msec) as long as the target and distractor objects are the same size or spatial frequency (unfiltered, low bandpassed, or high bandpassed). Such sequences were termed homogeneous. When targets (which were present in only half the sequences) were of a different size or scale than the distractors (heterogeneous sequences), a marked advantage (pop out) was observed for large (unfiltered) and low spatial frequency targets against small (unfiltered) and high spatial frequency distractors, respectively, and a marked decrement for the complementary condition, as in the example shown in Fig. 13. Importantly, this pattern of results for large and small images was unaffected by holding absolute (Expt. 7) or relative (Expt. 8) spatial frequency content constant over the different sizes, and it could not be explained by simple luminance- or contrast-based pattern masking. These results suggest that size/scale tuning in object recognition was accomplished over the first several images (< 576 msec) in the sequence (within which a target never appeared), and that the size tuning, contrary to expectations of the Scaling Hypothesis, was implemented by a mechanism sensitive to spatial extent rather than to variations in spatial frequency.



**Fig. 11.** The stimuli used in the different experiments. Size changes (maximum extent of the object, in degrees) are represented on the horizontal axis, changes in spatial frequency filtering (center frequency, in cycles per degree) on the vertical axis. The first single presentation naming experiment, the Pure Size RSVP, and the two masking RSVP experiments (Expts. 1, 3, 5, 6) used unfiltered large and small images. The second single presentation verification experiment and the Pure Scale RSVP experiments (Expts. 2, 4) used large size images filtered at two center frequencies with a 1.5 octave wide bandwidth. The Absolute RSVP experiment (Expt. 7) used large and small images filtered around 10 cpd, whereas the Relative RSVP experiment (Expt. 8) used different center frequencies in proportion with the size changes between large and small images. The only untested condition, small images filtered around 2 cpd, would create unidentifiable blobs.



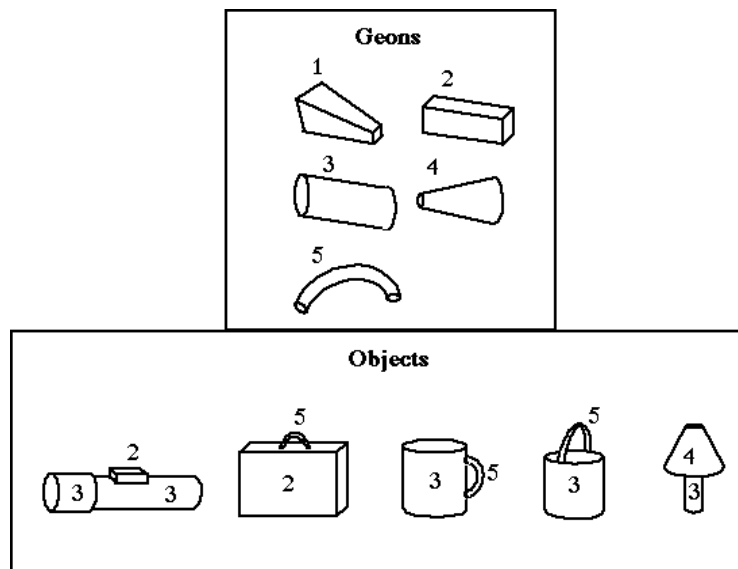
**Fig. 12.** A heterogeneous trial example from the Pure Scale RSVP experiment (Expt. 4 in Fig. 11). In a sequence of high bandpassed images, a low bandpassed image, which might or might not be the target, appeared in half of the heterogeneous trials. All images in all conditions had the same size. Each image was shown for 72 msec. (The arrow represents time).



**Fig. 13.** A heterogeneous trial example from the Pure Size RSVP experiment (Expt. 3). In a sequence of small images there is one large image in a random position which might or might not be the target. In the other type of heterogeneous sequence, one image would be small and all the others large. The arrow represents the time axis (there were no visible frames around the images). In other experiments, the target and distractors could vary in spatial frequency.

## Recognition of Objects in the Presence of Rotations in Depth: Learning from Human Recognition

Another challenge to implementing a vision system is how to achieve the invariance that people demonstrate in recognizing objects at different orientations in depth (up to parts occlusion). Biederman [Biederman, 2000] reviewed the literature and concluded that many of the phenomena of object classification can be derived from a viewpoint-invariant representation of an object's parts (geons) and relations, termed a geon structural description (GSD) (Fig. 14). The viewpoint invariance derives from a specification of the orientation and depth discontinuities in an image surface in terms of differences in properties that are invariant over orientation, such as whether an edge is straight or curved, pairs of edges parallel or not, or the kind of vertices that are formed from the cotermination of edges. Such a representation: (a) enables the facile recognition of depth-rotated objects, even when they are novel, (b) provides the information that is employed not only to distinguish basic-level but also highly similar members of subordinate-level classes, and (c) enables mapping onto verbal and object-reasoning structures. This work has defined a critical set of subgoals in adapting the Gabor jet model: (a) extraction of the orientation and depth discontinuities, (b) characterization of these discontinuities in terms of viewpoint invariant properties, and (c) segmentation of the object from its background and the different parts of an object from each other.



**Fig. 14.** Five geons (from a vocabulary of  $< 50$ ) and five objects. Note that the pail and the bucket are composed of the same geons but in different relations. TOP-OF is a defined spatial relationship. If the page is rotated  $180^\circ$ , the pail will resemble a cap and the lamp a trowel or shovel.

## Visual Representations: Analysis of Tradeoffs and Feature Learning

In order to recognize a diverse set of objects in realistic environments, a vision system will need to use a variety of features in its decision making process. We describe work below on an object recognition model that incorporates features similar to those believed to be used in mammalian vision systems, and includes among these features a set based on Gabor wavelets. A key feature of this model, besides its ability to recognize diverse sets of objects with high accuracy, is that it is readily parallelizable, and has the potential of relatively efficient mapping onto the photonic multichip module architecture described herein.

During the research grant period, as explained below, an effort was undertaken to analyze the tradeoffs that apply to visual representations, based on combinations of detected features like those of the human visual system. Bartlett Mel's SEEMORE model [Mel, 1997] and its extensions involve taking combinations (*e.g.*, conjunctions and disjunctions) of detected features (*e.g.*, edges, line segments, and blobs), and is inspired by similar operations in the human visual system. Parameters traded off in this investigation included the number of objects in a scene, the amount of clutter, the object complexity, and the number of features and combinations of features needed for accurate recognition. This work provides direction toward developing an efficient processing model and mid-level representation for the optoelectronic eye hardware.

The visual object recognition system in the brain appears to be organized in a hierarchy within which cells at higher levels respond to more complicated combinations of features than cells at lower levels, and simultaneously are more invariant to translation, rotation, and distortions. An architecture that builds both feature complexity and invariance gradually is evidently capable of solving the enormously difficult problems of invariant object recognition in a real world environment. Similar architectures have been proposed for machine vision [*e.g.*, Fukushima, *et al.*, 1983], and have had limited success in some domains. We believe that such systems can potentially be extended to solve some of the much more difficult problems that biological systems can – if a better way of designing the feature detectors is found.

To this end, we studied several design tradeoffs governing visual representations based on learned, spatially-invariant conjunctive feature detectors, with an emphasis on buffering such systems against false-positive recognition errors – von der Malsburg's classical “binding” problem. We then derived an analytical model that makes explicit how recognition performance is affected by the number of objects that must be distinguished, the number of features included in the representation, the visual complexity of individual objects, and the clutter load, *i.e.*, the amount of visual material in the field of view in which multiple objects must be simultaneously recognized, independent of pose, and without explicit segmentation.

Using a complex artificial visual domain to model object recognition in cluttered scenes, we have shown that this analytical model achieves good fits to

measured recognition rates in simulations involving a wide range of clutter loads, object complexities, and feature counts. We have further developed a novel “greedy” algorithm for feature learning, derived from the analytical model, which grows a representation by choosing those conjunctive features that are most likely to distinguish objects from the cluttered backgrounds in which they are embedded. We have shown that the representations produced by this algorithm are compact, decorrelated, and heavily weighted toward features of low conjunctive order. Our results provide a more quantitative basis for understanding when spatially invariant conjunctive features can support unambiguous perception in multi-object scenes, and lead to several insights regarding the properties of visual representations optimized for specific recognition tasks.

These results further suggest not only optimized combinations of features, and hence feature detectors, for efficient mapping onto the emerging hybrid electronic/photonic hardware platform, but also a potential generalizable algorithm for developing such optimized combinations based on either assumptions or empirical data regarding both the objects to be detected, and the backgrounds in which they are likely to be embedded.

### **Extraction of Shape-Defining Contours**

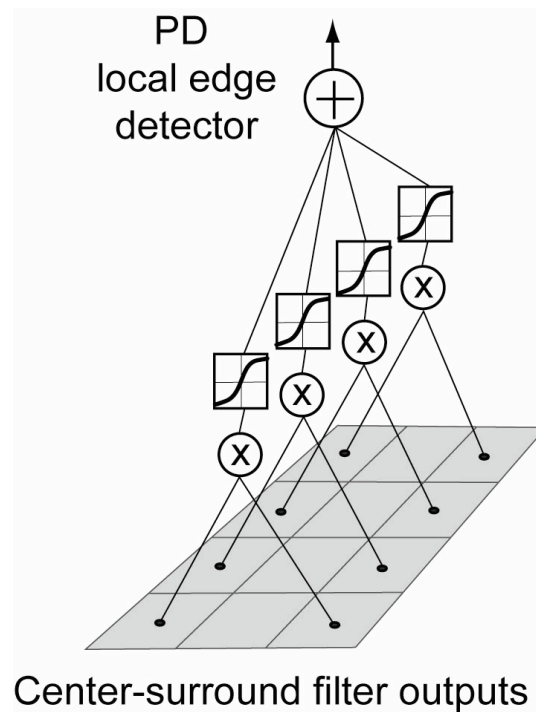
Automatic recognition of objects in visual scenes is critically needed if machines are to locate, identify, and manipulate objects autonomously in the human world. In collaboration with our ARO-sponsored MURI partners at USC, one of our principal long-range goals has been to develop neuromorphic algorithms that transform image sequences into a contour-based format – expressing information similar to that contained in an artist’s line drawings (see, for example, Figs. 18 and 19 in the section below on the contour extraction network). Visual representations of this kind are extremely data-compressed, but remain intelligible to human observers [Biederman and Ju, 1988]. Realtime hardware systems capable of automatically generating line-drawing-like contour representations of images will have many applications, include intelligent remote sensing/surveillance, visual control of autonomous mobile agents, and optical object/scene recognition.

Over the course of the entire funding period, we have taken two significant steps towards the development of hardware-mappable algorithms for visual contour-extraction. First, we developed a high-performance local oriented edge detector called the pairwise-difference-of-gaussian (PD) detector, with excellent edge-analyzing properties and an underlying nonlinear structure designed to be directly mappable onto a layered optoelectronic substrate. Second, we developed a neuromorphic contour-extraction network that takes local oriented PD edge filters as inputs and produces line-drawing-like representations at the output.

## Extraction of Shape-Defining Contours: The PD Edge Detector

We developed a local oriented edge operator, called the Pairwise-DOG (PD) detector, that combines the outputs of several pairwise products of center-surround filters straddling a candidate edge (Fig. 15). Though based exclusively on local intensity cues, the PD operator has proven to be a highly effective oriented edge detector, emphasizing those contours that are most related to object identification, and outperforming the venerated Canny detector (Fig. 16). Furthermore, most images acquired outside a well-lit office environment exhibit a large dynamic range of intensity values, and thus contain both very strong and very weak edges and contours. In such cases, the PD detector exhibits a distinct advantage over other oriented linear (*e.g.*, Gabor functions, as used by [Lades, *et al.*, 1993]) or nonlinear filters [Iverson and Zucker, 1995], in that the edge-detected images are far less sensitive to variations in the detection threshold.

Examples of PD filter outputs applied to images are shown in Fig. 16, in which a black pixel indicates that a PD filter of any orientation exceeded the display threshold at that pixel.



**Fig. 15.** Pairwise-difference (PD) edges were computed as follows: At each of 8 neighboring locations along the edge axis (only 4 are shown), a pair of center-surround filter outputs is multiplied and passed through a sigmoidal nonlinearity, and then summed. PD values were computed at 8 orientations at each pixel.



## Extraction of Shape-Defining Contours: The Contour-Extraction Network

A line drawing, including the main occluding boundaries, orientation discontinuities, and pigment contours, contains most of the shape information needed to recognize objects and scenes (Biederman & Ju 1988). The primate visual cortex is a high-performance image processor designed in part to solve this problem through long-range, multi-scale, dynamical interactions needed for contour completion and grouping [Grossberg and Mignolla, 1985; Peterhans and von der Heydt, 1989; and Kapadia, *et al.*, 1995]. The non-classical surrounds of V1 receptive fields appear to be the first stage in the visual-cortical stream to carry out these computations. For example, contour elements that are aligned with the orientation of the classical receptive field of a neuron can facilitate the cell's response. This effect likely depends on the extensive network of long-range horizontal connections among cortical cells.

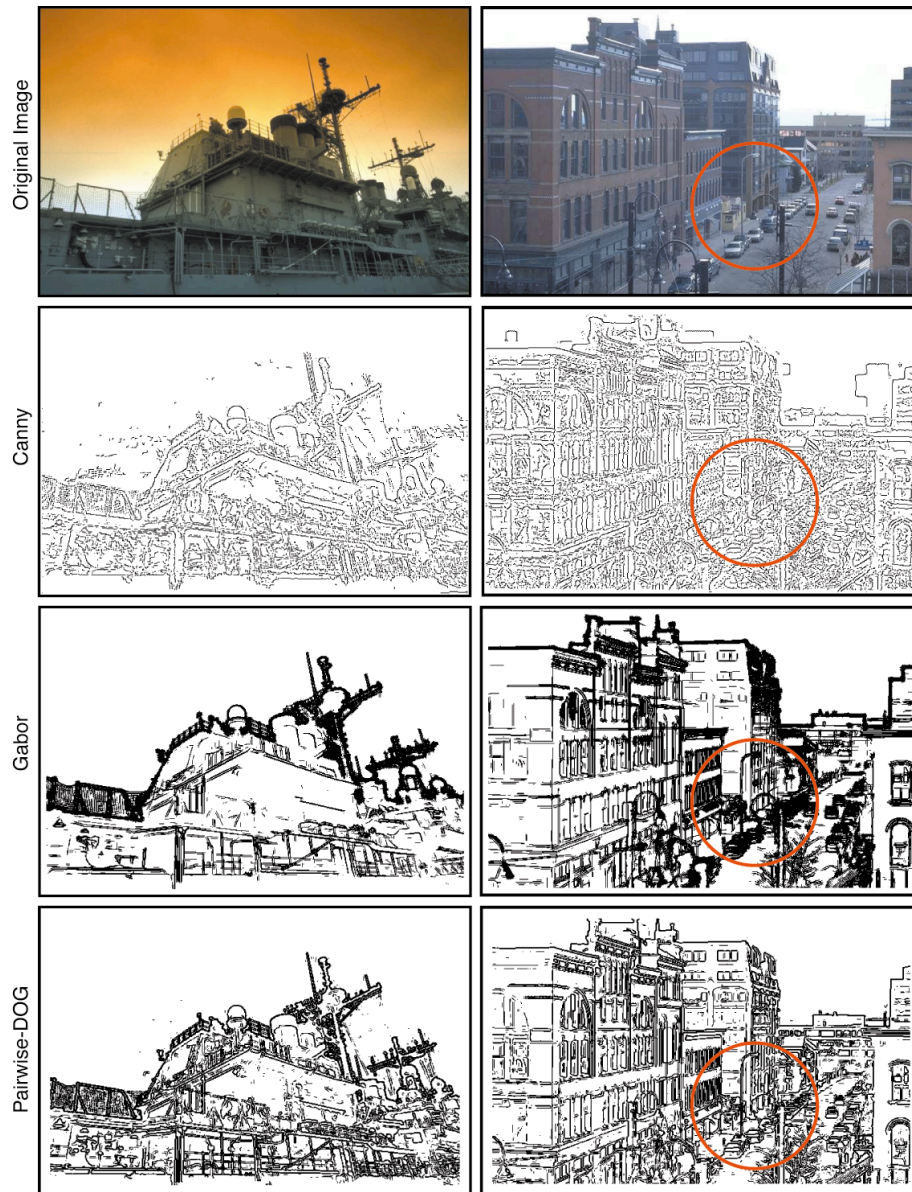
To address the problem of contour extraction, we developed a neuromorphic contour extraction network as shown schematically in Fig. 17. The model was implemented as follows: Let  $g_i$  represent the probability that a contour exists at location  $(x, y)$  with orientation  $q$  and scale  $s$  (see Fig. 17). Our approach assumes that the influences that determine  $g_i$  can be expressed in terms of 5 quantities present in, or derived from, the  $N$ -dimensional input vector  $x = \{x_1, \dots, x_N\}$ , in which the input components  $x_i$  represent local PD edge filters (as shown in Figs. 15 and 16). The network schematic shows the input features represented at the bottom and the contour-hypothesis output units at the top. Four derived quantities appear in between the input units and the outputs. The objective of this massively parallel network is to arrive at a valid binary contour interpretation for the whole image in a small number of iterations.

Though the network is dynamical, we assume that its job is to reach a stable contour interpretation quickly. Here, we have no concern for the state-space trajectory of the network. We therefore formulate its update rule in terms of the steady-state relationship between each contour hypothesis  $g_i$  and the five influences that determine its value:

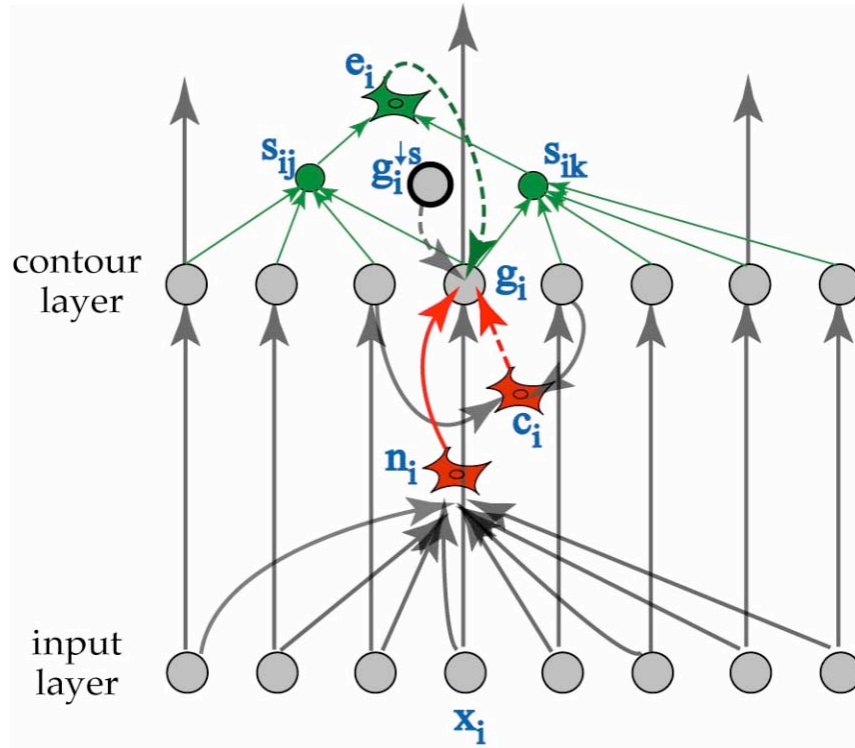
$$g_i = \text{nonmax}_t \left( \text{modul} \left( \frac{\alpha_t x_i + \beta e_i}{1 + \gamma_t n_i^2}, g_i^{\downarrow s} \right), c_i \right)$$

in which the terms are defined as follows:

- $x_i$  is the response of the local-edge unit in the input layer corresponding in position, orientation, and scale to contour hypothesis  $g_i$ .
- $a_t$  is a time-dependent coefficient that depresses the influence of  $x_i$  through time.



**Fig. 16.** Comparison of original images and outputs of three low-level edge detectors; circles facilitate detailed comparisons of corresponding regions of images. First row: Original images. Second row: Canny edge results include hysteretic thresholding. Third row: A linear sine-wave Gabor filter produces somewhat better results, especially for elongated contours. Gabor filters tend to produce thick lines at high contrast edges, and often miss lower contrast edges which are perceptually salient. We found that small changes in the threshold setting led to large changes in the resulting edge-detected image—a serious drawback when automated processing is required. Fourth row: Results from our PD filter. Output is much less sensitive to the absolute value of the final threshold, leading to good representation of fine details over entire dynamic range of edge contrasts. The PD filter provided the input to our contour extraction network (see Fig. 17).

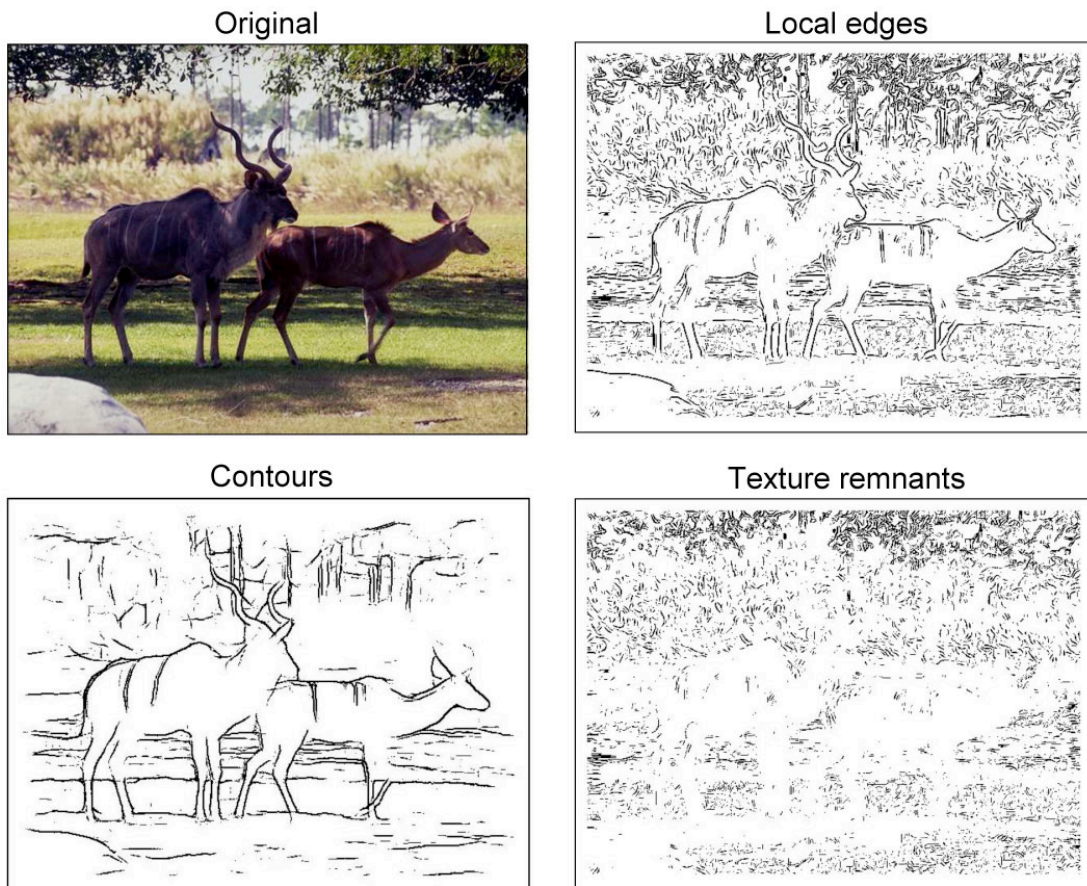


**Fig. 17.** Schematic illustration of the contour detection network. The input layer contains a large bank of local oriented edge features  $x_i$  corresponding to PD filter outputs. The output layer contains contour hypothesis units  $g_i$ . Dashed lines denote feedback pathways leading back to the  $g$  layer. Long-range contour shape subnetwork (in green) contains high-threshold sigmoidal subunits  $s_{ij}$  that receive input from surrounding contour hypothesis units; the overall long-range evidence signal  $e_i$  is computed as the MAX of all subunit responses. Inhibitory interneurons are shown in red. Feedforward inhibition acts divisively on  $g_i$  and is mediated by the interneuron labeled  $n_i$ . Feedback inhibition in the output layer represents the spatial mutual-exclusion among groups of incompatible contour hypotheses; non-maximum suppression is mediated by inhibitory interneuron  $c_i$ .

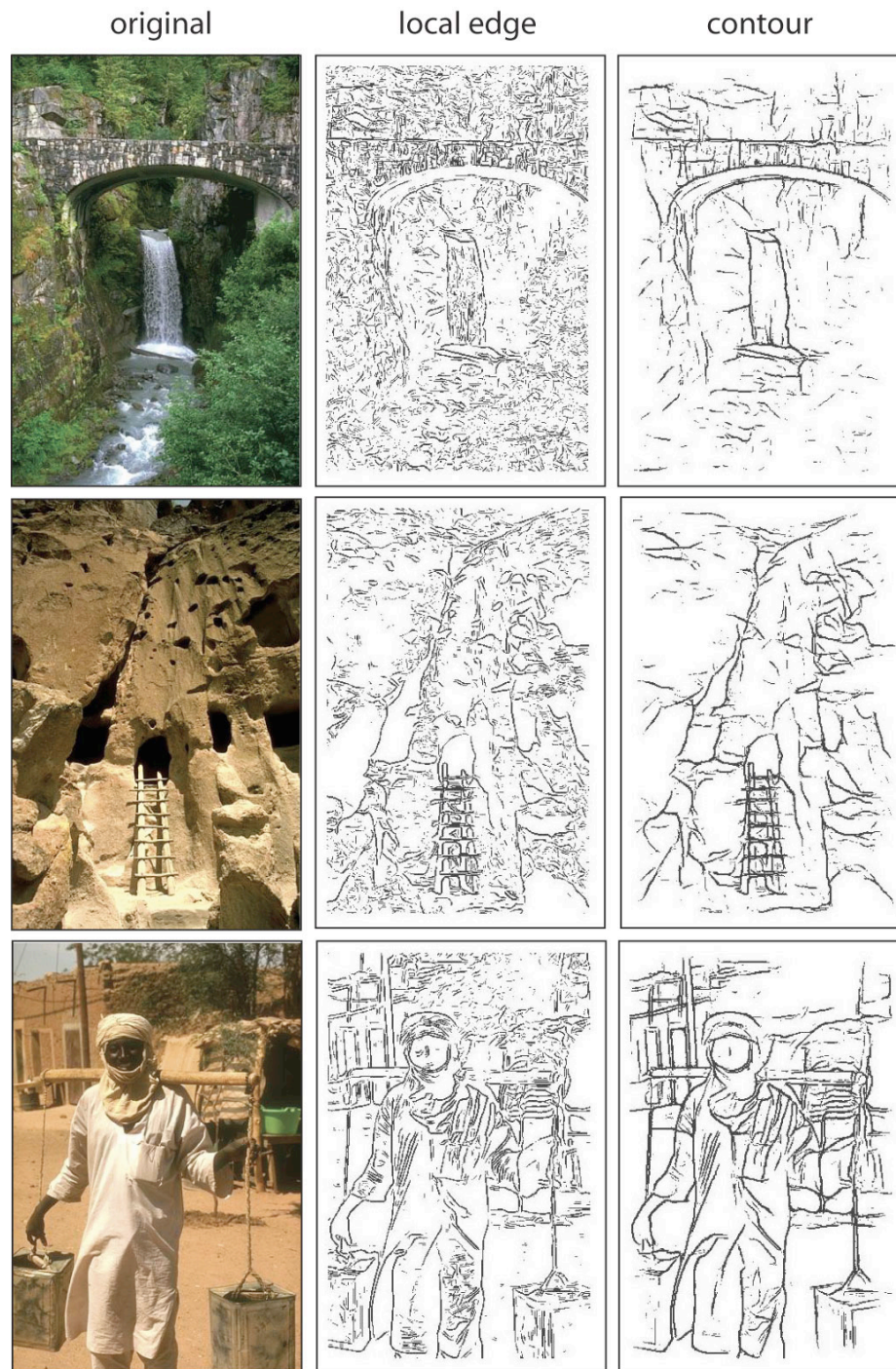
- $e_i$  is the output of the long-range contour detection sub-network, whose goal is to detect any well-formed contour that includes contour hypothesis  $g_i$ . The sub-network contains 250 contour prototypes derived from a training set of 30 silhouette images of common objects. Each prototype is represented by a high-threshold sigmoidal subunit  $s_{ij}(g)$  with 100 inputs from other contour units originating in the  $60 \times 60$  pixel region surrounding  $g_i$ . Evidence from the set of contour prototypes is combined using a max operation, namely,  $e_i = \max(s_{ij})$ . We use this disjunction-like form, because summing evidence over mutually incompatible prototypes is statistically nonsensical;  $b$  is a constant.
- $n_i$  provides a measure of the overall “edginess” of the surround of contour hypothesis  $g_i$ , and is thus monotonically related to the probability of false-positive contour recognition in that region of the image.  $n_i$  combines inputs from local edge units of all orientations in a  $60 \times 60$  surround. However,  $n_i$  contains an iso-orientation bias towards the orientation of  $g_i$ . Scaled by  $t$ ,  $n_i$  acts to normalize the strength of the positive contour evidence provided by  $e_i$ . The normalizing signal is feedforward, carried by inhibitory interneurons projecting from the input to the output layers.
- $g_i^{\wedge s}$  represents the corresponding contour hypothesis in the output layer but at a lower spatial resolution. This variable allows the incorporation of coarse-scale contour cues (e.g., from color or texture) into the network iteration.
- The function “modul,” in turn, represents our best guess from first principles as to the conditional probability of contour  $i$  given both high-resolution evidence (first term) and low-resolution evidence ( $g_i^{\wedge s}$ ). It shows that strong evidence for a contour at low-resolution, by itself, provides only weak evidence for any particular high-resolution contour hypothesis  $g_i$ . However, low-resolution evidence dramatically boosts the believability of even small quantities of high-resolution evidence.
- $c_i$  represents a second type of inhibitory influence that captures the notion of spatial mutual exclusivity among physically incompatible contours. Thus  $c_i = \max_{j \in C_i}(g_j)$  is the maximum response among  $g$  units in the “conflict” set  $C_i$  of  $g_i$ . In its simplest form (which we have so far implemented),  $C_i$  includes contours at the same position with very similar orientation. This set also includes contours of identical orientation at nearly the same position.  $C_i$  is thus strongly orientation tuned, and is communicated between incompatible contour hypotheses through lateral inhibitory connections in the output layer.
- “nonmax” implements non-maximum suppression with  $\text{nonmax}(a, b) = a$  if  $a > b$ , and 0 otherwise. This function mediates spatial-mutual-exclusion inhibition in the output layer, since it allows only one contour hypothesis to survive in each conflict set.



The results of applying the contour network to a complex scene are shown in Fig. 18. A dense local edge map is converted through a few network iterations into a sparse line-drawing-like representation. Subtracting the contours from the dense edge map shows local edges rejected by the network. Little shape information survives this operation, supporting the idea that shape information critical for human and machine object recognition is contained in the network output. Additional examples of network output are shown in Fig. 19.



**Fig. 18.** Local edges compared with long-range shape-defining contours. A. Original image. B. Local edges used to seed long-range contour network. C. Contours approximate line drawing of scene. D. Contours subtracted from local edge map leaves only amorphous texture remnants.



**Fig. 19.** More contour-extraction network results. From left to right: originals, local PD edge maps, and resulting contour images. Examples illustrate extremely dense texture-sensitive local edge maps and their conversion by the network to a sparse contour representation emphasizing object shapes.

## **Adaptive Fusion of Cues**

While many models and vision algorithms focus on one important aspect of the vision problem, such as those for detecting edges as discussed above, robust and general object recognition and scene segmentation stands to benefit greatly from employing a combination of cues. In fact, another lesson we can learn from biology is the necessity to flexibly use a rich array of such independent cues (including color, motion, stereo and many more) by combining them in a situation-dependent way. Only with adaptive cue fusion does it seem possible to reduce the enormous unreliability and ambiguity inherent in each single cue. We have proceeded here on two fronts. For one, we have developed specific hardware architectures to extract and fully exploit rich sensor signals. Reliable motion extraction, for instance, necessitates signals with high spatial and temporal resolution and very efficient motion extraction mechanisms. Secondly, we have developed techniques to adaptively combine different cues with each other in a situation-dependent way.

For example, target detection and tracking is difficult in realistic conditions due to the enormous variability of target properties, illumination conditions, trajectory characteristics, changes in background, and partial or temporary occlusion. Any one cue thus is highly unreliable. As an easily available sample domain we set ourselves the goal of detecting the head (the "target") of a person walking across the field of a video camera. As cues we have used skin color, pixel intensity change, contrast range, shape, and motion continuity. Cues all vote for head positions in the scene, and a consensus is created as a weighted average. Each cue computes its own confidence level on the basis of its agreement or lack thereof with the consensus. This confidence level is turned into dynamically changing relative cue weights. In addition, individual cues adapt their internal parameters so as to reach optimal agreement. Thus, the shape cue (which is initially totally blank) adapts to the gray level distribution around the consensus point, or the color cue shifts its concept of skin color. The only cues with an initial prejudice are skin color and pixel change. All other cues acquire their parameters only during a given trial. The system performs with remarkable reliability in the presence of active attempts to frustrate it by changing the color of illumination, by having the person change motion direction, by having another person cause occlusion, or by changing backgrounds in order to create a loss of contrast [Triesch and von der Malsburg, 2000].

## **Development of a Bayesian Vision Model from the Mapping Perspective**

We have developed a Bayesian vision model, of which the computation can be formulated such that it maps onto, and computes efficiently on, the photonic multichip module (PMCM) structure. This model can be briefly described as follows. To reflect the statistics of visual scenes, at the lowest level, the underlying image takes on a Markov Random Field (MRF) prior that promotes smoothness while allowing for discontinuities (*i.e.*, edges). The observed image is a noisy version of this underlying image. (Additional modules, such as the prior for contours and the prior for texture, can be factored in at succeeding



levels.) The smoothness is embodied by coupling between nearby pixels, which, if broken, would tolerate abrupt changes (Fig. 20).

The visual computation thus deals with inferring the underlying image—including its features such as contours and texture—given an observed image. As this MRF formulation is what is called a “stiff” problem, if one exclusively relies on iterative lateral propagation of data between adjacent pixels, as typically done in the literature, the convergence time is unfortunately proportional to  $L^2$  or  $L^4$  [Blake and Zisserman, 1987], depending on the specific model, in which  $L$  describes the spatial extent, measured in number of pixels, of data communication.

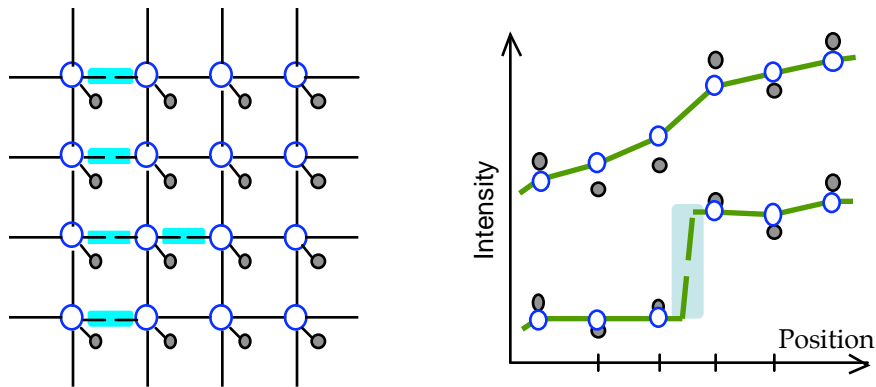
Inspired by principles of primate visual systems, we have developed an approach to reformulate the above visual inference problem into a non-iterative feedforward fan-in and fan-out computation [Huang and Jenkins, 2006], which can be readily mapped onto, and compute efficiently on, the PMCM structure. We have shown that the inferred image consists of two terms:  $x_0 = W^T y$ , and  $x_1 = W_1^T H W_1 y$ , in which  $y$  is the input image,  $W$  is a matrix representing the spatially-invariant fan-in weights,  $W_1$  is a subset of  $W$ , and  $H$  is a matrix pertaining to edges in the image (Fig. 21);  $x_0$  is an intermediate result that is the estimated hidden state (pixel value) resulting from a spatially invariant filter;  $x_1$  is the correction term which takes into account edge information. The resulting image estimate is  $x_0 + x_1$ .

The full matrices  $W$  and  $H$  are large ( $N^2 \times N^2$  and  $E \times E$ , respectively, in which  $N$  is the number of image pixels and  $E$  is the number of edge-detected pixels). By using their block-circulant properties and suitable approximation techniques, they can be represented by much smaller matrices while retaining excellent results. The resulting algorithm then can be implemented with reasonable hardware resources while still providing fast computation.

Examples of this model used for image de-noising are shown in Fig. 22. It is evident that better results are attained with the edge-preserving model (fourth row in Fig. 22) than with a more straight-forward space-invariant filter (third row in Fig. 22).

The above-described work demonstrates how a neurobiologically inspired vision algorithm can be re-expressed in terms of non-iterative feedforward fan-in and fan-out computations. Once the algorithm has been expressed in these terms, the tools described in previous reports can be used to map it onto the PMCM hardware architecture.

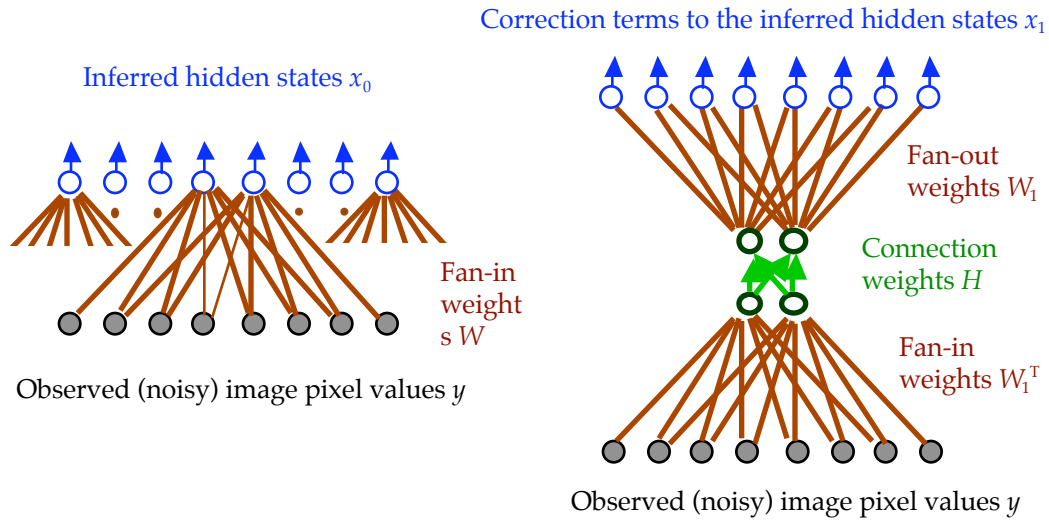




**Fig. 20.** Schematic diagram of the Markov Random Field (MRF) prior.

**Left:** A portion of the MRF. Solid and broken lines denote intact and broken coupling, respectively. Open circles denote hidden states and filled circles denote observed image pixels.

**Right:** Two examples of 1-D slices of the MRF, each based on the same intensity values of observed image pixels (filled circles). In the upper slice, all hidden states (open circles) are tightly coupled and therefore feature a smooth profile. In contrast, the lower slice contains a “broken” coupling (dashed line), across which the hidden states are no longer constrained to be of similar values. Consequently, the estimated hidden states can retain the sharp transition (shaded strip).



**Fig. 21.** Flow chart of probabilistic inference based on the MRF.

**Left:** The inference of hidden states (*i.e.*, restored image) without the presence of broken couplings in the MRF. This is a convolution, which can be implemented by space-invariant fan-in operations on our photonic multichip module (PMCM) structure.

**Right:** The computation of the correction terms – due to edges – to the inferred hidden states. This is achieved by passing through three layers of weights: (1) a set of fan-in weights; (2) connection weights  $H$  (in green); and (3) fan-out weights (which are identical in value to the fan-in weights). This feedforward flow of computation can also be readily implemented on our PMCM structure.



**Fig. 22.** (Previous Page) Image denoising results.

**Top row:** Original images.

**Second row:** Images corrupted with Laplacian (left column) and Gaussian (right column) noise, respectively. (Laplacian noise has a heavier tail than the Gaussian counterpart.)

**Third row:** Estimate of the original images given the noisy input images, disregarding edge information. These are produced by convolving the input images with a fan-in kernel that characterizes a bell-shaped weighting profile, which amounts to spatially invariant blurring of the input noisy image.

**Fourth row:** Estimate of the original images given the noisy input images, with considerations of possible abrupt changes in intensity levels between adjacent pixels. It can be understood as the results in the third row (*i.e.*, uniformly blurred versions of the noisy input images) plus correction terms aiming to restore the sharpness of edges. The correction terms can also be computed in a fan-in and fan-out style.

### **Investigation of Space-Time Tradeoffs in the Optoelectronic Eye Hardware System**

We have initiated the development of implementation modes that trade off hardware space and computation time in different ways. These modes are geared toward implementation of low- and mid-level vision algorithms. Because the human visual system, as compared with our projected optoelectronic hardware system, is 5 orders of magnitude larger in number of processing elements (neurons), approximately the same factor higher in number of interconnections, and approximately 6 orders of magnitude slower in response time, the understanding and utilization of techniques that can trade off space for time and *vice versa* is crucial for optoelectronic-hardware implementations of robust neurally inspired vision models.

One technique for trading off space and time, which we have begun to investigate, is to employ various degrees of programmability in each pixel. The primate visual system consists of a hierarchy of 2-dimensional layers of processors (neurons), in which each processor appears to represent a specific visual feature, determined by its fixed connectivity (on short time scales) to the feature detectors in antecedent layers from which it receives input. Given that neurons in the human visual system are extremely numerous (order of 10 billion), so that a large number of processors are available to analyze each pixel, this hardwired, or “labeled line”, approach is not particularly limiting.

As mentioned above, however, the architectures that we are considering will likely have room for significantly fewer processors than the biological counterpart, typically one processor or less per image pixel in a given layer, but will operate at clock rates that are orders of magnitude higher. As such, we have focused on ways to take advantage of the time dimension more effectively by

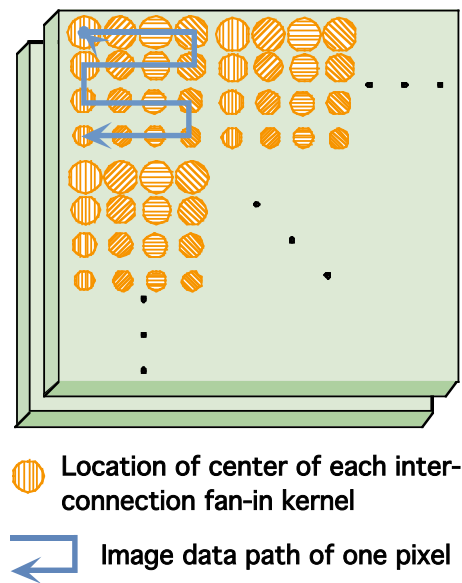
introducing programmability into the architecture, so that a given site in each array of pixel-locked processors may compute different features in a time-multiplexed fashion. The inclusion of such programmability also affords a more general processing architecture that can be differently configured for different visual domains or tasks. This work has led to the development of specific new multiplexing techniques, as described below.

### **Development of Spatio-Temporal Multiplexing Techniques**

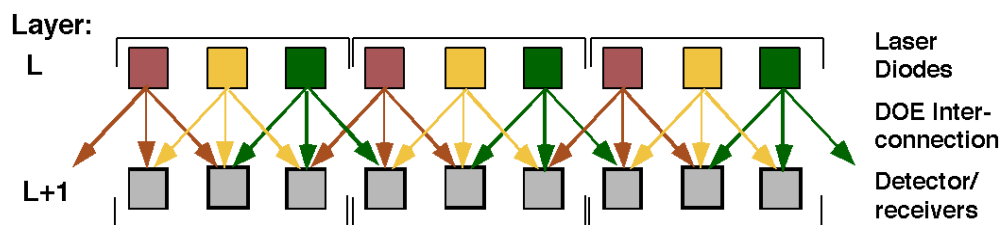
We have devised two spatio-temporal multiplexing techniques for implementation on the photonic multichip module hardware. Both techniques allow us to exploit the time dimension more fully than would a more direct mapping without such multiplexing. We note that in typical operation scenarios of the photonic multichip module, frame rates of the input images will be of order 10 to 1000 Hz, and bandwidths on each interconnection line between subsequent pairs of layers will be much higher, of order 10 to 100 MHz. Thus, for these multiplexing techniques to be computationally useful, the computation and layer-to-layer interconnection patterns should effectively change many times for a given input image.

We have devised a scrolling technique that can provide this feature; it uses fixed optical interconnections that implement various filters (or fan-out patterns) over space. The input image data is repeatedly shifted (laterally, in the plane) in time; at any one time-step, the image data is sent through the filters in parallel. At each subsequent time-step, a different portion of the image is centered over each filter, and thus a different set of computations is performed. Viewed from the subsequent layer, each fan-in pattern corresponds to a given filter over a different portion of the image (Fig. 23); the filtered image data in this plane is typically shifted in synchrony with the image data in the previous plane. Viewed from the point of view of the image data, a parallel set of various filters is being shifted across the image in time. An example of the application of this technique is to a full 4-D Gabor transform, in which different transform kernels are laid out in space over the array. Over time, each region of the image data is filtered by each different kernel.

Our second technique relies on partial sequencing of the output ports of the processing elements in time. It will be illustrated here for a 1-D array that employs laser diodes as the output port elements; the principle also applies to a 2-D array and to modulator-based systems. We will describe this for a system that has space-invariant layer-to-layer interconnections, with fan-outs of  $M$ , and physical (DOE) interconnection weights that have been fabricated to be unity. If used in a mode that provides complete programmability of the layer-to-layer interconnection weights, every  $M^{\text{th}}$  laser diode is turned on in the  $L^{\text{th}}$  layer. In the  $(L+1)^{\text{th}}$  layer, the gain of each receiver circuit is set proportional to the desired weight of the optical interconnection line that is currently active. In Fig. 24, for example, every third box (laser diode output) of layer  $L$  and all of the interconnections from those boxes are the same shade of gray, indicating that they are all on (active) during the same time step. During this time step, the three different weights for the three lines fanned out from a given laser diode are



**Fig. 23.** Spatiotemporal multiplexing by image data scrolling. Different interconnection patterns are laid out in space. Over time, the input and output image data is shifted; at each shift position, the image data is sent through the interconnection patterns (filters) to the next layer. Each pattern shown represents the location of the center of each interconnection fan-in kernel. (The spatial extent, or receptive field, of each interconnection fan-in kernel is not shown, and covers a number of such locations.) The lateral path of one pixel of image data is shown as a solid line and arrow.



**Fig. 24.** Spatiotemporal multiplexing by temporal modulation of processing-element inputs and outputs; shown for a 1-D array with an optical fan-out of 3. In one time step, every third laser diode and its DOE interconnection fan-out pattern are activated (shown in a common grayscale level). Gains of the detector/receiver circuits are set proportional to the desired interconnection weights. An interconnection round is completed in 3 time steps.

implemented as electronic gains in the three receiving circuits. In this example it takes three time steps to sequence through all the outputs of layer  $L$ , which completes one interconnection round. The sequence of three signals received at each receiving circuit is summed (integrated) over the time of one interconnection round. More generally, for a fanout of  $M$ , an interconnection round is completed in at most  $M$  time steps. By varying the detector circuit gains in time, multiple interconnection filters are effectively multiplexed in time at each spatial location. An example of the application of this technique is to a full 4-D Gabor transform in which different kernels are implemented in different time steps, and the image data is spatially stationary in time.

### **Enumeration of Issues for Mapping of Vision Algorithms onto the Optoelectronic Multichip Module Hardware**

As a guiding principle for our work on the mapping of vision algorithms onto the optoelectronic multichip module hardware, we have postulated various scenarios for the mapping of specific algorithms, namely von der Malsburg's dynamic link architecture and Mel's object recognition system. At a lower level, we have also been investigating the mapping of specific operations, such as linear convolution operations, multiple parallel convolutions with different kernels, and nonlinear on-center off-surround filters. This work brought out many issues and questions that were addressed by the algorithm subgroup and the hardware subgroup. Issues included the need for space-time tradeoffs (mentioned above); the degree to which programmability is appropriate; analog, digital, and hybrid number representations; storage and shifting of analog and digital signals; and tradeoffs between in-the-plane electronic interconnections and plane-to-plane optical interconnections.

### **Development of a Nature/Nurture Algorithm for Visual Adaptation**

An adaptive vision sensor placed in the real environment must be capable of adapting to changes in the environment such as lighting conditions, view aspect, pose, and many others if it is to provide robust object recognition. Adding adaptive capability to the Photonic Multichip Module (PMCM) architecture requires an innovative strategy, as the envisioned hardware implementations contain two interpenetrating sets of interconnections. The feedforward (out of the plane) weights supplied by the diffractive optical elements are fixed following an initial (off-line) training period, and adaptive weights are conveniently added by implementing lateral connections within each plane by means of additional interconnecting VLSI circuitry.

During the research program, a novel Nature/Nurture (N/N) algorithm for calculating the weight sets in such an architecture (consisting essentially of two interpenetrating neural networks) was proposed and examined. The PMCM hardware implementation and its correspondent interpenetrating neural network model are shown in (Fig. 25). The modified multilayer perceptron (MLP) model can map onto the PMCM structure in the feedforward data processing scenario after all of the feedforward weight values are determined. The detectors in each artificial neuron collect the optical fan-in signals to perform



a summation function. The collected information is then transformed by a sigmoid transfer function circuit and transmitted to the associated DOE layer. The vertical weighted fan-outs in the MLP model represent the optical weighted fan-outs from the DOE. The lateral weighted interconnections in the MLP model represent the lateral VLSI interconnections modulated by adaptive weight memory devices in each VLSI chip, interconnecting nearest neighbor and perhaps next-nearest neighbor pixels.

The general learning steps for the Nature/Nurture algorithm are as follows: (1) In an MLP structure such as the PMCM architecture, the interlayer (vertical) weights are first trained by using an original dataset and a supervised learning algorithm, with the lateral weights disabled (set to an initial value). This step constitutes the "Nature" procedure. (2) The values of all of the vertical weights in the MLP are then fixed. (3) The original training dataset is then modified in some manner, for example by adding a background or noise to produce a new dataset. (4) The hidden layer (lateral) weights are then turned on, and trained by using the new dataset. This step constitutes the "Nurture" procedure. (5) The final performance of the MLP is then examined against both the original and modified data sets to determine if adaptation has increased the robustness of object detection.

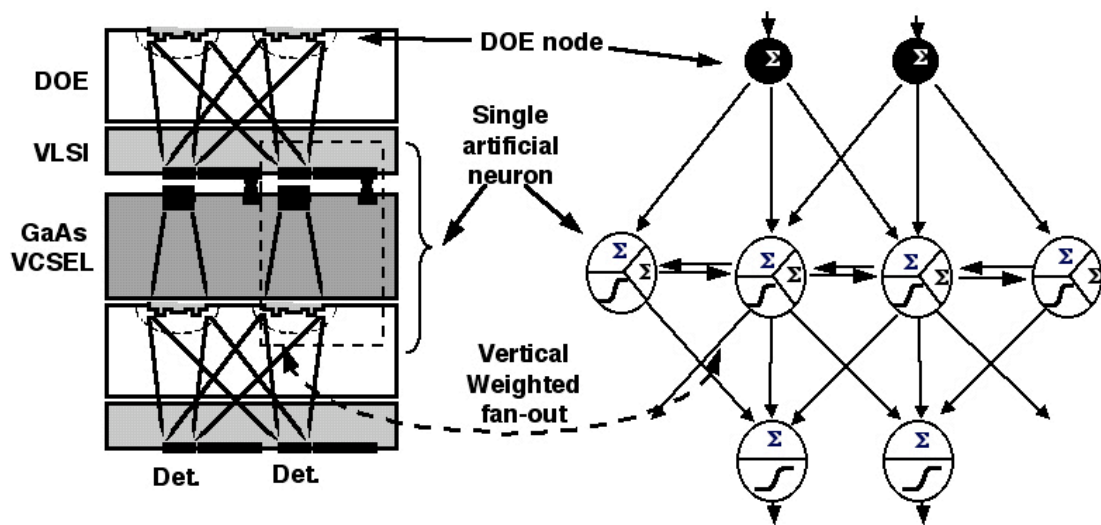
In the PMCM hardware implementation, the DOE patterns representing the vertical weight values are fixed and established after the fabrication process (Nature). On the other hand, the adaptive electronic analog memory devices representing the lateral weight values can be adapted after fabrication (Nurture). The traditional error backpropagation (EBP) algorithm is suitable for finding the vertical weights only. An analytical formula for calculating the lateral weights was then derived based on the EBP algorithm. A universal Nature/Nurture algorithm for finding vertical and lateral weights for an arbitrary layer MLP was also derived.

An optical character recognition (OCR) task was simulated by using an artificial 100-100-2 MLP (Fig. 26). The Nature/Nurture algorithm was adopted to train the 100-100-2 MLP by means of a MatLab program on a Macintosh G3 PowerPC. This 100-100-2 MLP with local interconnections models the signal pathways in the PMCM architecture. The 10 by 10 array of nodes (unity gain detectors/emitters) in the first layer are matched to a corresponding 10 by 10 array of DOEs in the first layer. The second layer's 10 by 10 neuron unit array represents an array of 10 by 10 artificial neurons. Each artificial neuron is assumed to consist of a dual-rail VLSI neuron unit, a VCSEL driven by the VLSI unit, and a DOE unit. The adaptive weighted local lateral interconnections between artificial neurons are also modeled in this algorithm. A final 2 by 1 VLSI neuron array is placed beneath the DOE-VLSI-VCSEL-DOE module to collect the ensemble of the weighted optical outputs. The final two neurons produce two individual output values. These two indices can be compared with the desired target values for classifying the input patterns.

A dataset of 10 by 10 digital pixellated versions of the characters "1", "2", "3" and "4" was initially prepared (Fig. 27). Another nine datasets produced



from the original dataset with different degrees of gray level gradient background were also prepared (Fig. 28). The 100-100-2 MLP was first trained in the Nature procedure by using the original dataset. After the Nature procedure alone, the MLP was able to fully recognize the original dataset, but not the modified datasets (Fig. 29(a)). After the lateral weights were modified by means of the Nurture training procedure operating on the modified datasets, the MLP was capable of recognizing both the new modified dataset and the unmodified (original) dataset with 100% recognition accuracy, as shown in Fig. 29(b).



**Fig. 25.** Mapping a single hidden layer feedforward neural network (right) onto the stacked PMCM architecture (left). The vertical weighted fan-outs represent the optical weighted fan-outs from the DOEs. The lateral weighted interconnections represent the lateral links modulated by the adaptive weighted memory devices in the VLSI plane. The detectors in each artificial neuron collect the optical fan-in signals to perform a summation function. The collected information is then transformed through a sigmoid transfer function and transmitted to the next DOE layer. Both the PMCM and the model are extendable sideways. Here only sections of the architectures are illustrated for simplicity.

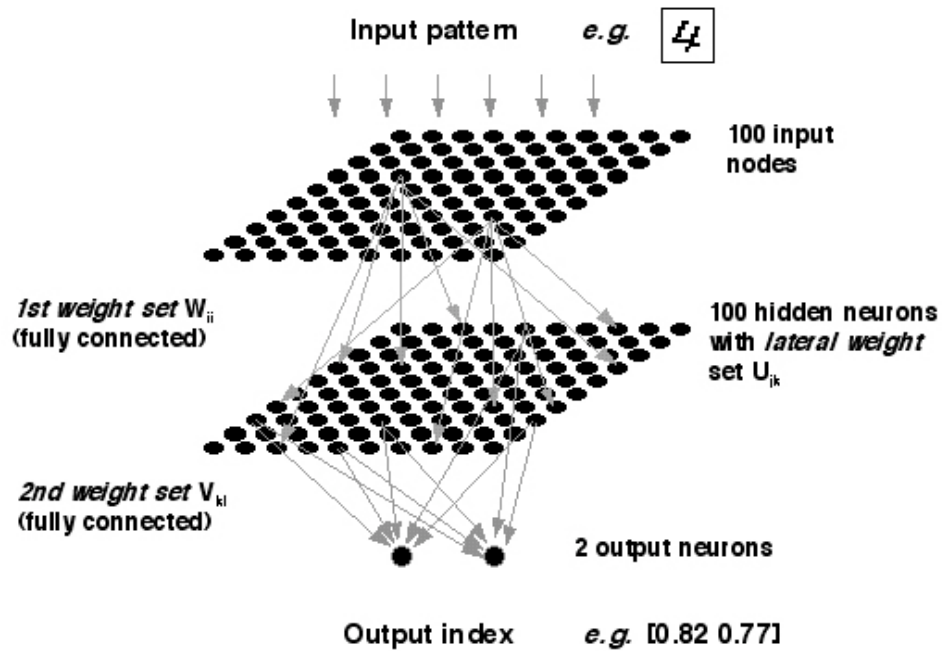


Fig. 26. A 100-100-2 perceptron with vertical and lateral weight sets.

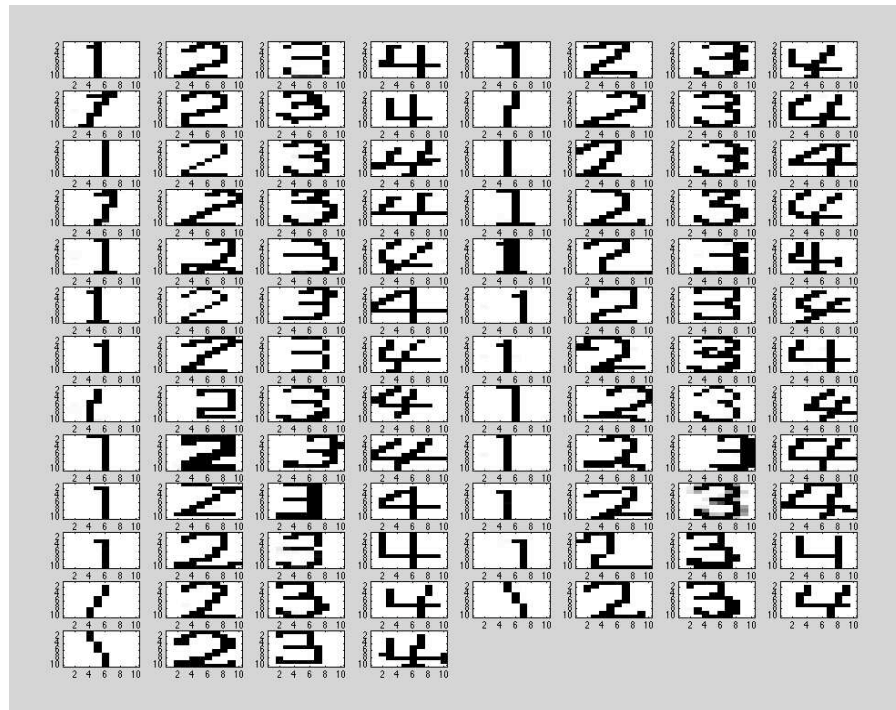
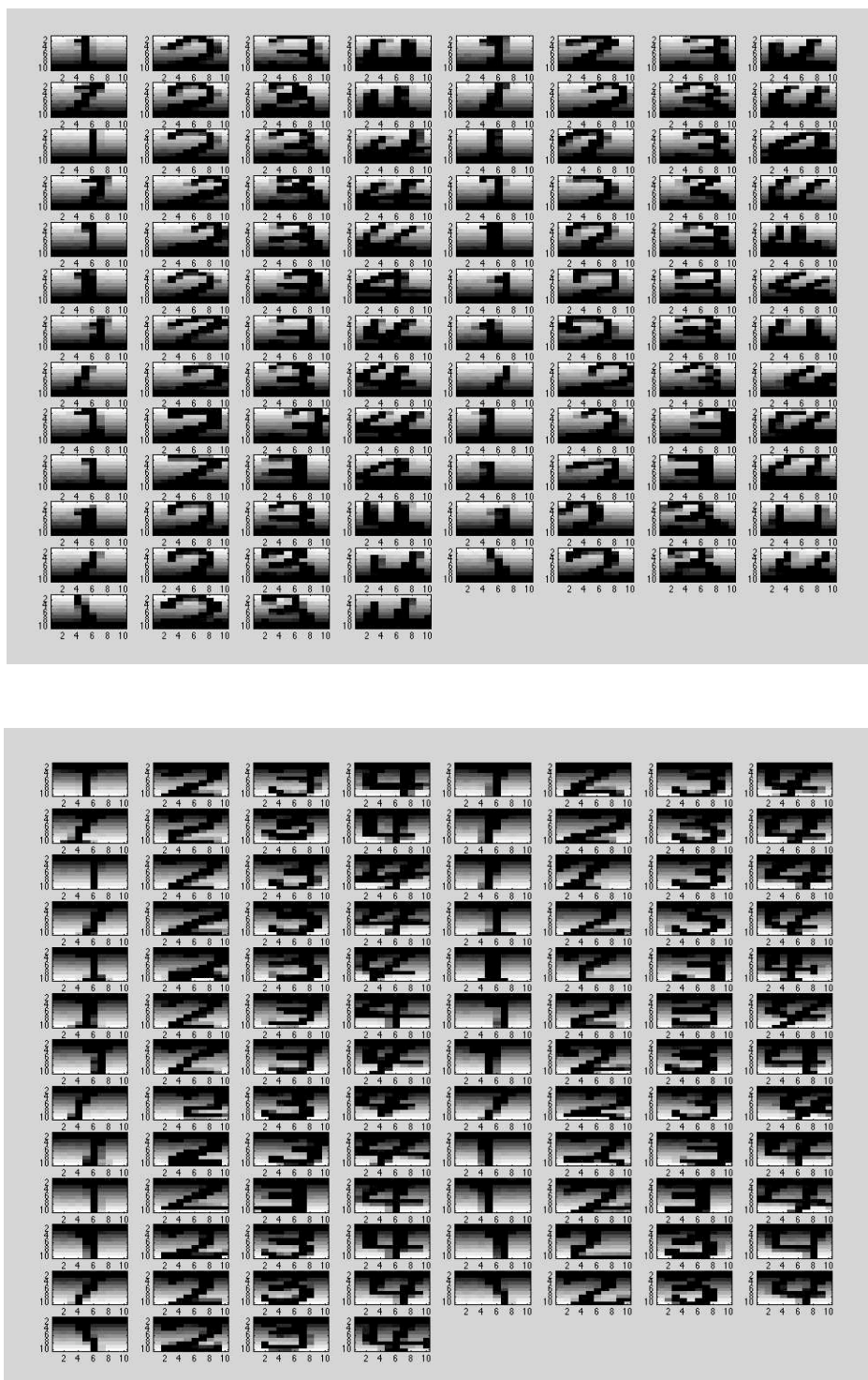
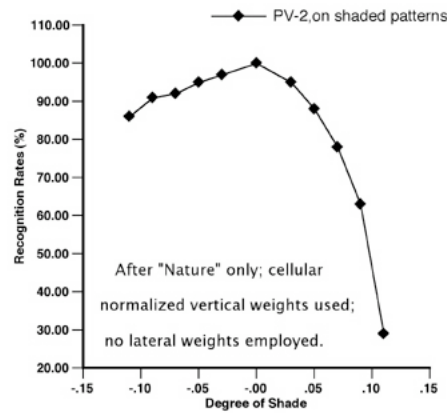


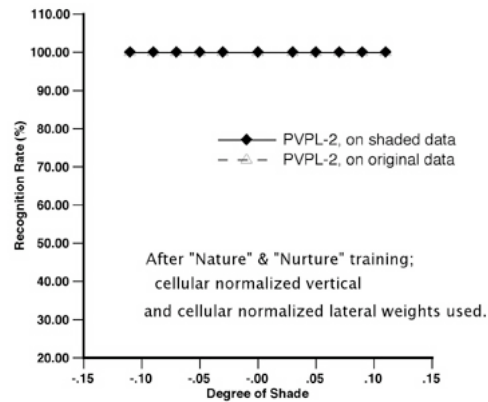
Fig. 27. Images of the original dataset for the “Nature” stage of the procedure in the novel Nature/Nurture algorithm.



**Fig. 28.** Sample images of modified datasets for the “Nurture” stage of the procedure. The degree of shading in the upper dataset is  $-0.11$ . The degree of shading in the lower dataset is  $+0.11$ .



(a)



(b)

**Fig. 29.** (a) Recognition rates for the shaded datasets as a function of the degree of shading of the background added to the original dataset after the “Nature” stage only. The recognition capability for the shaded dataset drops dramatically as the degree of shading increases. In this case, the as-trained vertical weights from the original (unshaded) dataset are used during the “Nature” stage of the procedure. (b) Recognition rates for the shaded datasets and the original dataset as a function of the shaded degree of the background after the “Nature” and “Nurture” stages. The recognition rates for all datasets are now all 100%. In this scenario, the normalized, the original vertical weights are used during the “Nature” stage of the procedure; both the original vertical weights and as-trained lateral weights are used in the “Nurture” stage of the procedure.

## **Significant Accomplishments: Hybrid Electronic/Photonic Hardware Implementation**

One of the key issues in the proposed hybrid packaging approach to the implementation of adaptive optoelectronic eyes is the successful demonstration of high-density fan-out/fan-in interconnections among the various layers of the hybrid MCM stack. This in turn involves demonstration of 2-D arrays of silicon VLSI photodetection/local processor units, 2-D arrays of MQW modulators or VCSELs, the interconnection of the silicon and gallium arsenide active element arrays by flip-chip bonding, the design and fabrication of appropriate high density fan-out/fan-in optical interconnections using diffractive optical elements (DOEs), the incorporation of focal power either within the DOE design itself or by means of a separate microlens array, the antireflection coating of all of the optical interfaces within the stack, the incorporation of absorbing material ("optical black") wherever reflections or unwanted diffracted orders cannot be tolerated, and the packaging of the multiple layers into an integrated functional block including attention to both alignment issues and thermal dissipation concerns.

Given the funding constraints imposed at the outset of the grant, it should be noted here that the emphasis of the funded research component of this MURI effort was on vision algorithms, models, and architectures. In this section, we describe the significant accomplishments achieved during the grant period in the area of hybrid electronic/photonic hardware implementation, which proceeded as a secondary focus leveraged by other resources. As funding under the related DARPA Photonic Wavelength and Spatial Signal Processing (PWASSP) initiative became available ("Dense 3-D Integrated Photonic Multichip Modules for Adaptive Spatial and Spectral Image Processing Applications"; Start Date: 15 June, 2000), support for the hardware implementation component was greatly increased and resulted in significant additional leverage to the MURI program.

### **Evaluation of Dual-Input, Dual-Output Silicon VLSI Neuron Unit Arrays**

During the initial stages of this multi-year research program, while the research effort on optimal parsing of the functionalities to be included in each layer of a multichip module implementation of an adaptive optoelectronic eye was still in progress, it might have appeared premature to define and implement specific functional units in silicon VLSI chips. On the other hand, many of the issues inherent in developing a functional hybrid electronic/photonic hardware implementation are to first order independent of the nature of such specific functional units. Therefore, during the research program we undertook (in parallel with the algorithm, vision model, and architecture effort) to examine the scientific and technological roadblocks to eventual hybrid MCM implementation based on readily available and/or modifiable hardware components.

One of several possible silicon chip designs for eventual incorporation in an adaptive optoelectronic eye involves the combination of photodetection capability with analog, nonlinear transformations. The implementation of dense

fan-out/fan-in interconnections among paired processor layers is also highly suggestive of a neural network approach.

Under separate sponsorship, we had previously designed, simulated, and fabricated  $16 \times 16$  arrays of dual-input, dual-output silicon VLSI neuron unit arrays that implement parallel nonlinear functionality. In this case, the nonlinear function implemented is that of a sigmoidal transformation of the difference between the excitatory and inhibitory channel inputs, generating two separate outputs (for positive and negative differences of the inputs), a function that is in turn characteristic of numerous neural network models [Jenkins and Tanguay, 1992].

The arrays comprised  $100 \times 100$   $\mu\text{m}$  pixels, fabricated in the  $1.2$   $\mu\text{m}$  MOSIS CMOS process, within which were arranged two photodetectors (vertical junction photodiodes) representing positive (excitatory) and negative (inhibitory) channel inputs, dual current mirrors and nonlinear transformation stages, and dual output amplifiers that terminated in pads designed for eventual flip chip bonding to MQW modulator elements or vertical cavity surface emitting lasers (VCSELs) disposed on a separate III-V compound semiconductor substrate (representing positive (excitatory) and negative (inhibitory) channel outputs).

During the research program, we completed an evaluation of these existing  $16 \times 16$  neuron unit arrays, with a view toward determining their applicability for functional incorporation in the emerging hardware platform, as well as their capability for flip chip bonding to 2-D arrays of MQW modulators or VCSELs. Detailed measurements focused on the uniformity of response across the array, the functionality of the nonlinear transformation imposed, the bandwidth over which the nonlinear transformation can be effected, the power dissipation on a pixel-by-pixel basis, and the design flexibility that might potentially be afforded by redesign in the  $0.85$   $\mu\text{m}$  MOSIS CMOS process.

A key result from this investigation was the measurement of the output voltage as a function of input photocurrent (generated by an  $850$  nm laser diode co-integrated with the probe station employed for these measurements). The input-output transformation is indeed sigmoidal, with the data closely matching the SPICE simulation at low frequencies ( $< 10$  kHz). At the two higher frequencies measured ( $100$  kHz and  $1$  MHz), the functional dependence remained sigmoidal, but the saturation voltage decreased somewhat. Post-measurement analysis indicated that this frequency loading effect was due to the inadvertent absence of several key vias designed to couple  $V_{\text{dd}}$  supply lines fabricated in both metal layer 1 and metal layer 2, thereby reducing the current carrying capacity of these lines by a factor of two while simultaneously loading the drive lines with a large parallel capacitance. This deficiency could easily be corrected by refabricating the chip with an improved design.

The (pair of) photodiodes integrated within each pixel on the chip were designed for increased collection efficiency by using substrate rather than well collection. The measured responsivity at  $850$  nm for these photodiodes was  $0.254$  A/W, consistent with what has been reported by other investigators for



similarly designed structures. The use of bulk photogenerated charge collection, however, also increased the degree of crosstalk, roughly 8% for photodetectors separated by a single pixel spacing (100  $\mu\text{m}$ ), even though a  $p^+$  guard ring was included in this design in an attempt to minimize cross-pixel signal contamination. Although this crosstalk figure reduces the effective signal-to-noise ratio, most artificial neural network structures are robust enough to minimize the deleterious effects of this degree of pixel-to-pixel interaction. Subsequent designs will address the crosstalk issue while attempting to maintain relatively high collection (quantum) efficiency.

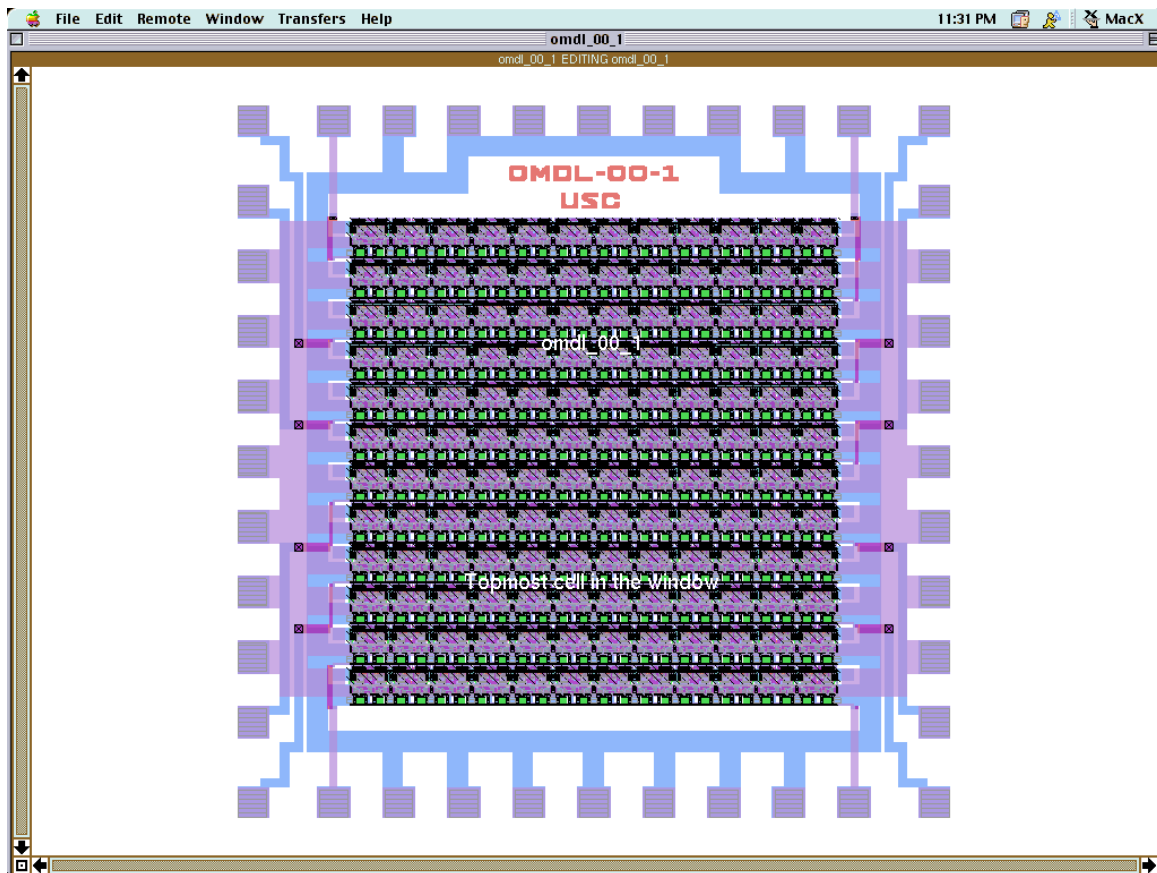
Since these chips were originally designed to accommodate MQW modulators (which require high voltages, but low currents), during the research program we evaluated the implications of redesigning the output driver circuits to accommodate vertical cavity surface emitting lasers (which instead require lower voltages and much higher currents). This evaluation proved that the existing chips could in fact be redesigned to accommodate VCSEL drivers without compromising the current pixel pitch significantly (moving from 100  $\mu\text{m}$  to 125  $\mu\text{m}$ , which matches the pitch of the VCSEL array masks that we designed for USC fabrication, as well as the masks used by our collaborators at the Army Research Laboratory (ARL; Dr. George Simonis and colleagues). In these modified designs, the VCSEL driver circuits occupied by far the greatest fraction of the chip real estate, an issue of concern for eventual downsizing of the array pitch. This issue places even more importance than before on the availability of low threshold, high efficiency vertical cavity surface emitting lasers, as described in more detail below.

In the process of designing for incorporation of Si CMOS VCSEL driver circuits, four new silicon VLSI chips were conceived, designed, and fabricated in total. Chip OMDL-00-1 includes a  $12 \times 12$  array of neuron units that incorporate 3 mA maximum-drive-current VCSEL drivers, two in each pixel, as shown schematically in the layout diagram below (Fig. 30). This chip was designed for direct indium flip chip bonding to correspondingly pixellated VCSEL arrays.

A variant of this chip, Chip OMDL-00-1-x, is functionally identical except for the fact that all of the bias lines are interconnected to provide a single common bias point for the entire array (thereby reducing the number of control lines from 14 to 3).

Chip OMDL-00-2 was designated as a companion test chip, and includes several key VCSEL driver components and subcomponents arranged for functional tests, as well as an array of the photodiodes included within each pixel, herein configured as a network interconnected by a hexagonal resistive grid.

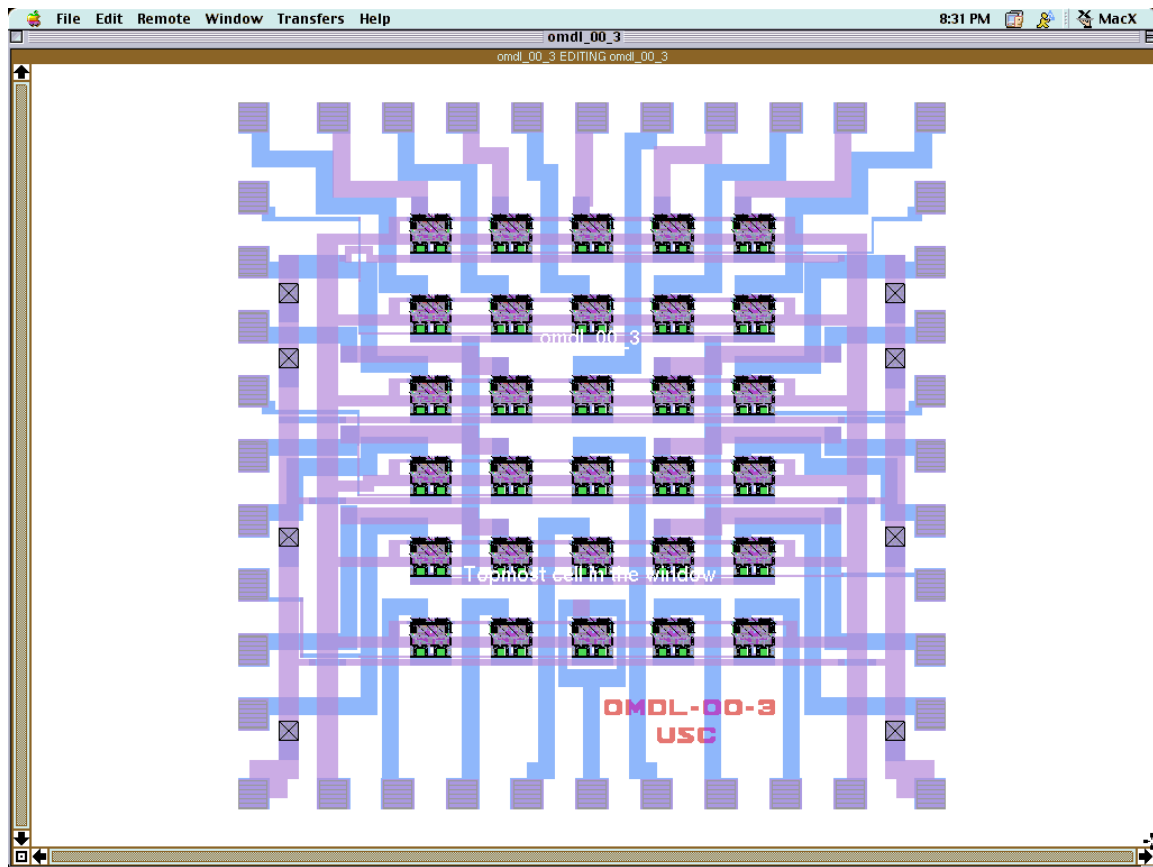
Chip OMDL-00-3 includes a  $5 \times 6$  array of neuron units with incorporated Si CMOS drivers, laid out for external wire-bonding to incommensurate (different pitch) VCSEL array elements, as shown in the layout diagram below (Fig. 31).



**Fig. 30.** VLSI layout of Chip OMDL-00-1, which comprises a  $12 \times 12$  array of neuron units that in turn incorporate 3 mA maximum-drive-current VCSEL drivers, two in each pixel, configured for direct flip-chip bonding to a mating GaAs VCSEL array.

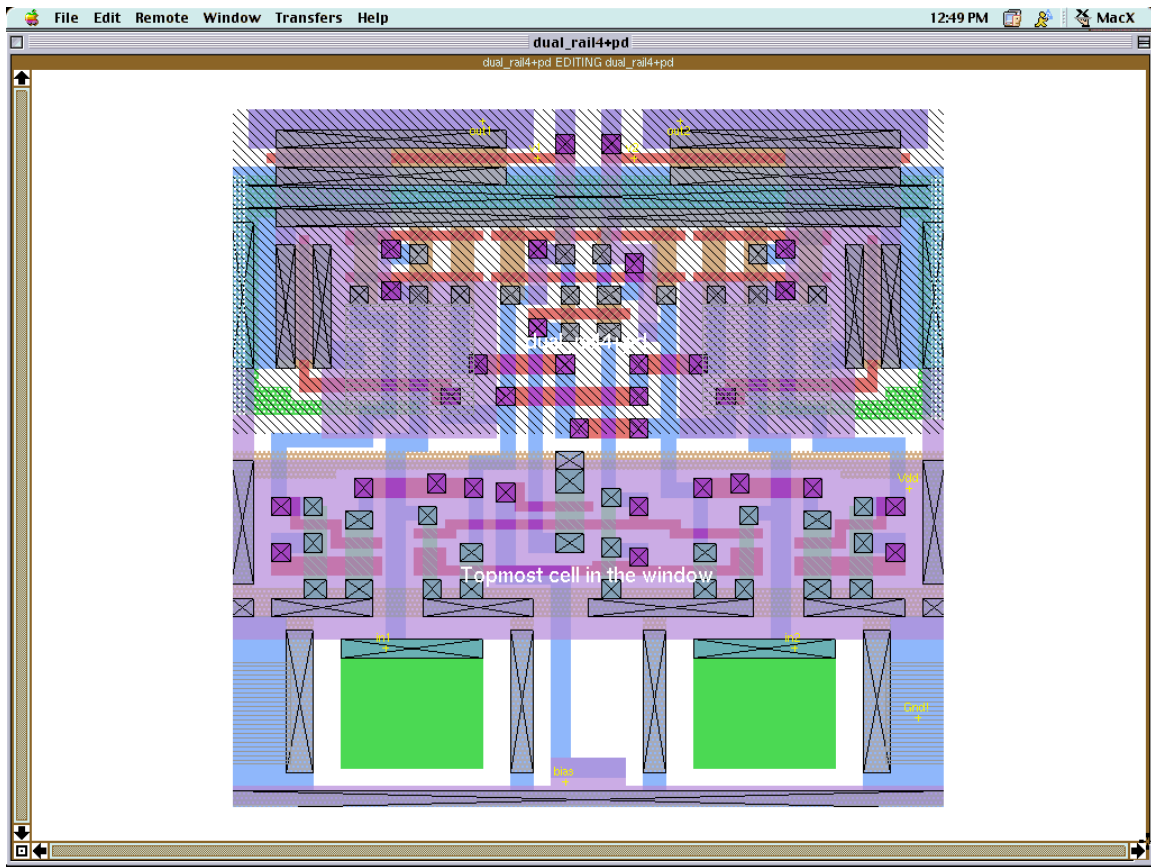


The layout of a single  $122.4 \times 124.8 \mu\text{m}$  pixel within Chip OMDL-00-1 is shown in the layout diagram presented in Fig. 32, and as shown schematically in Fig. 33. The dual rail construction of each pixel is evident, with each input signal starting from an *n*-diffusion/*p*-substrate photodiode (PD1 or PD2), followed by a linear-I-to-sigmoidal-V transformation circuit (M1XX or M2XX families). The final output stage comprises a pair of large PMOS transistors that are configured as operational transconductance amplifiers (OTA; M301 or M302). A bias circuit is included for adjusting the operating point of the entire dual-rail driver (M1 to M6).



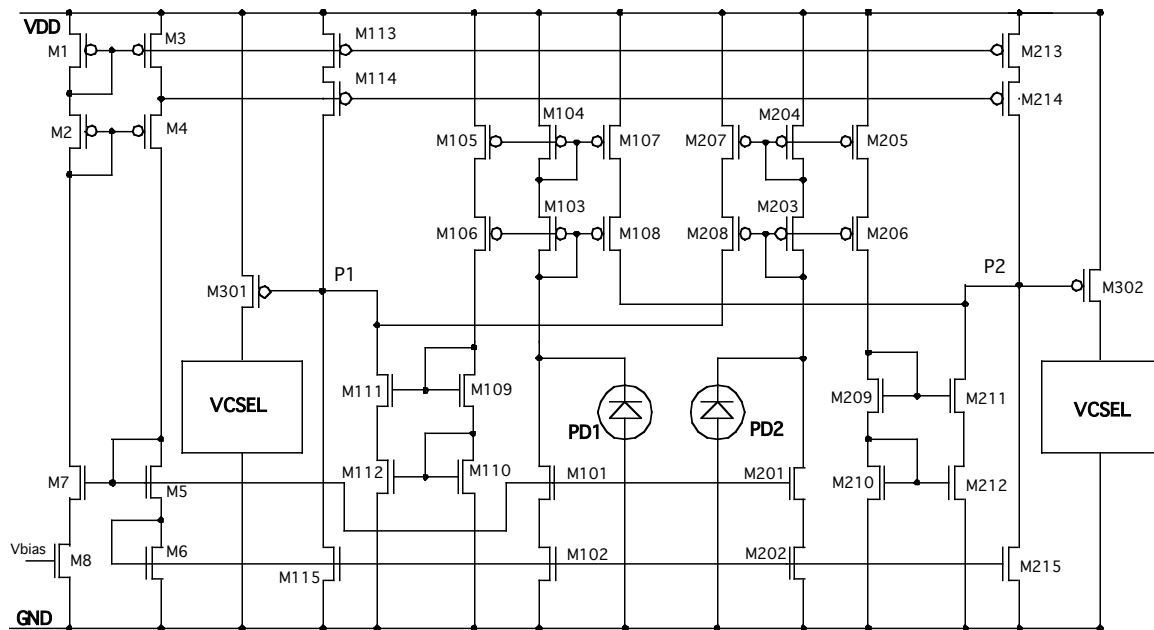
**Fig. 31.** VLSI layout of Chip OMDL-00-3, which comprises a  $5 \times 6$  array of neuron units that incorporate 3 mA maximum-drive-current VCSEL drivers, two in each pixel, configured for external wire bonding to a complementary GaAs VCSEL array.

Successful fabrication and testing of these dual-input, dual-output sigmoidal response neuron unit array circuits (including high current PMOS VCSEL drive transistors) was achieved during the research program period. For example, as-fabricated photographs of Chip OMDL-00-1 are shown below in Figs. 34 and 35. The 12-by-12 neuron unit array of this chip has a 125  $\mu\text{m}$  pitch as shown in Fig. 34, and with external wire bonding pads fit into a 2.2 mm by 2.2 mm chip area as shown in Fig. 35. The supply voltage was designed to be 5 V to reduce power consumption. The power dissipation was measured to be 13 mW per pixel, or about 1.87 W per chip in full operational mode at a frequency of 2 MHz. The large signal sigmoidal response was demonstrated at frequencies up to 1.5 MHz without distortion of the desired sigmoidal characteristics, as shown in Fig. 36. According to HSPICE simulations, the 3-dB cut-off frequency can be increased to about 15 MHz if the chip is fabricated by using the 3.3 V, 0.25  $\mu\text{m}$  TSMC process. The total neuron unit area could be reduced to about one fifth of the current size by using this same 0.25  $\mu\text{m}$  TSMC process, as shown in Fig. 37.

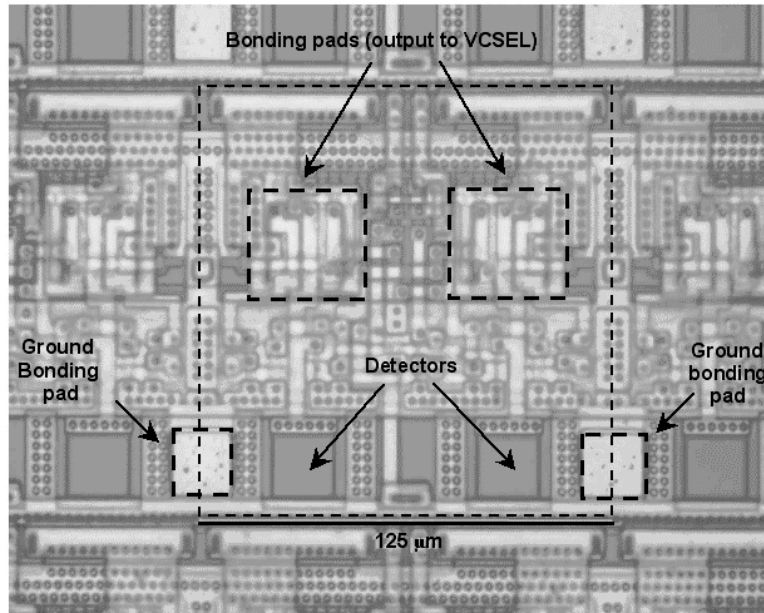


**Fig. 32.** VLSI layout of a single pixel within the  $12 \times 12$  array included in Chip OMDL-00-1, as shown in Fig. 30.

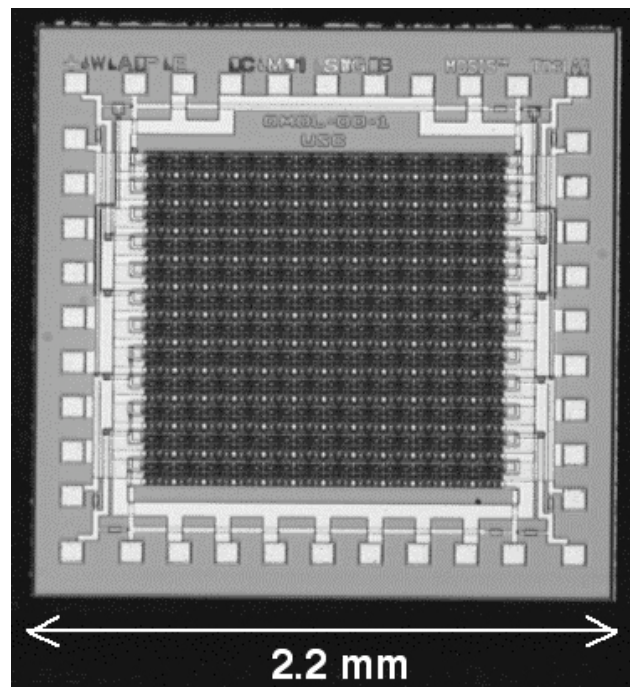
The successful fabrication and testing of these chips also allowed for direct flip-chip bonding of silicon neural unit arrays to VCSEL arrays, thereby enabling the evaluation of thermal management and heat dissipation, as well as direct coupling to complementary DOE arrays. We plan to include detailed test results of optically addressed Si CMOS chips driving externally mounted VCSELs within an array in the related DARPA/ARO PWASSP Final Progress Report. In addition, as low threshold VCSEL arrays become available that match the pitch of the currently fabricated Si CMOS chip sets, we will continue with flip chip bonding experiments to produce two-dimensional arrays of Si CMOS VCSEL drivers and VCSEL output devices that can be characterized for use in conjunction with DOE arrays, as described in the section on photonic multichip module (PMCM) integration below.



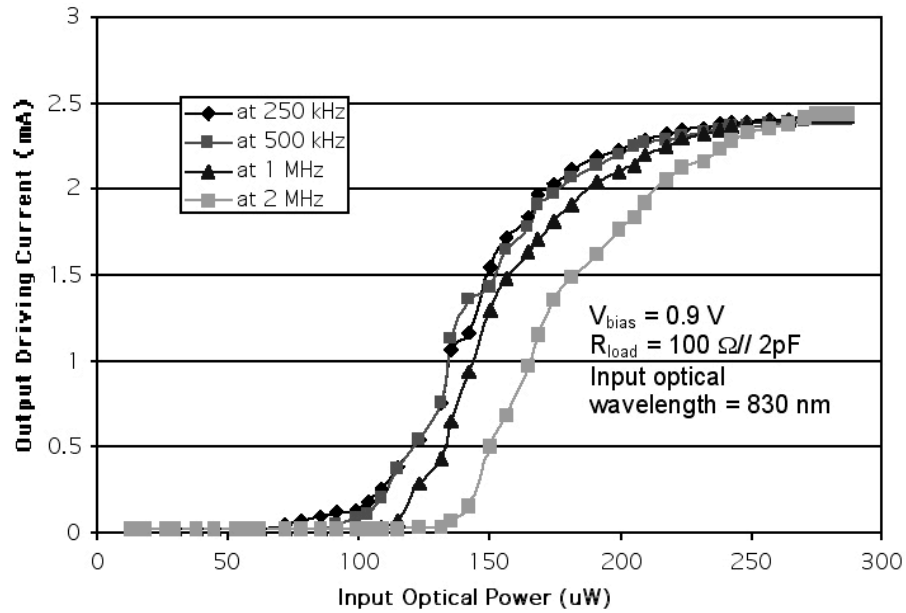
**Fig. 33.** Schematic diagram of the dual-input, dual-output sigmoidal response neuron unit array circuit, including high current PMOS VCSEL drive transistors, as described in the text.



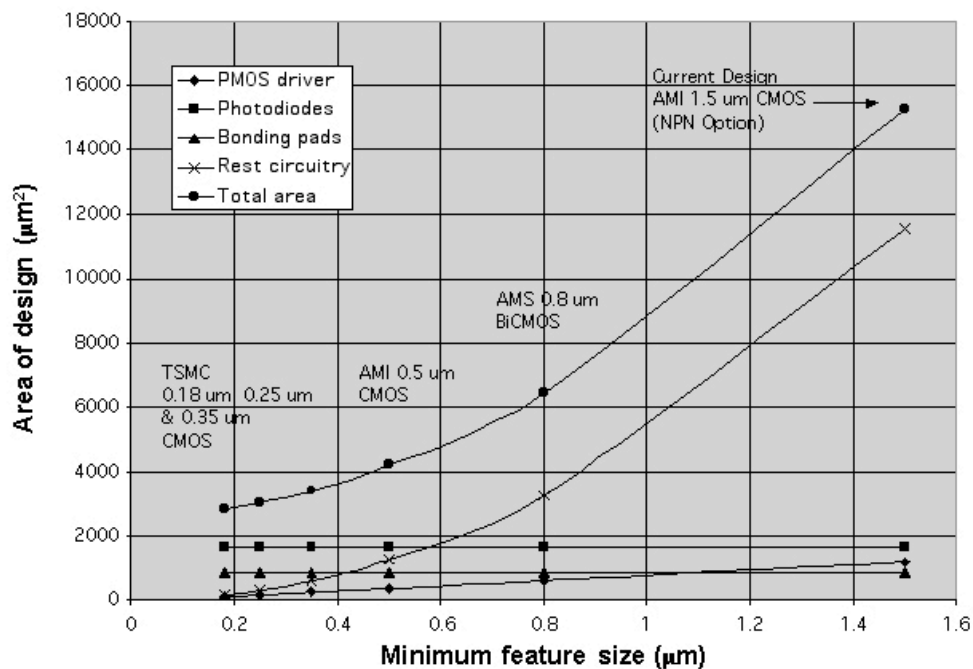
**Fig. 34.** Optical micrograph of a single pixel of the dual-input, dual-output sigmoidal response neuron unit array circuit (Chip OMDL-00-1), including high current PMOS VCSEL drive transistors, as described in the text.



**Fig. 35.** Optical micrograph of the dual-input, dual-output sigmoidal response neuron unit array circuit (Chip OMDL-00-1), including high current PMOS VCSEL drive transistors, as described in the text.



**Fig. 36.** Output current as a function of input optical power for the dual-input, dual-output sigmoidal response neuron unit array circuit (Chip OMDL-00-1), including high current PMOS VCSEL drive transistors, at several frequencies.



**Fig. 37.** Design area of key individual components as a function of minimum feature size for dual-input, dual-output sigmoidal response neuron unit array circuits including high current PMOS VCSEL drive transistors, as fabricated in several different VLSI technologies.

## Development of a Single-Sided Flip-Chip Bonding Process

During the research program, we undertook an aggressive investigation of advanced packaging technologies for integrating combinations of silicon-based and gallium-arsenide-based VLSI electronic and photonic chips. As can be seen in Figs. 1, 2, and 3, the interfaces between paired silicon photodetector/functional implementation chips and gallium arsenide 2-D modulator or VCSEL arrays are configured as face-to-face proximity couplings, with very high density vertical interconnections between pairs of corresponding elements. For the majority of the configurations that we currently envision, the interconnection density is in the range of 3 to 4 interconnections per pixel. For pixel sizes between  $100\text{ }\mu\text{m} \times 100\text{ }\mu\text{m}$  and  $50\text{ }\mu\text{m} \times 50\text{ }\mu\text{m}$ , the interconnection density range is therefore  $3 \times 10^4\text{ cm}^{-2}$  to  $1.6 \times 10^5\text{ cm}^{-2}$ .

A crucial element in the development of such an interface is thus the necessity for high-density parallel electrical interconnection of two-dimensional pad arrays on the silicon chips with corresponding two-dimensional pad arrays on the gallium arsenide chips, using flip-chip bonding approaches. In our laboratory, indium bump bonding by means of 10 to 30 micron sized indium bumps has proven to be highly reliable, and was investigated as the interconnection method of choice for this application. We have extensive experience with this indium bump flip-chip bonding approach, and have developed a novel indium "velcro" deposition process that allows micro-interpenetration of the roughened surfaces of two opposing indium bumps deposited on the mating surfaces, resulting in low impedance (few  $\Omega$ ), reliable, and temperature-compliant contacts as shown in Fig. 38. We have extensively tested arrays of these types of indium bump bonds, as shown in Fig. 39, over the temperature range of 2 K to 300 K.

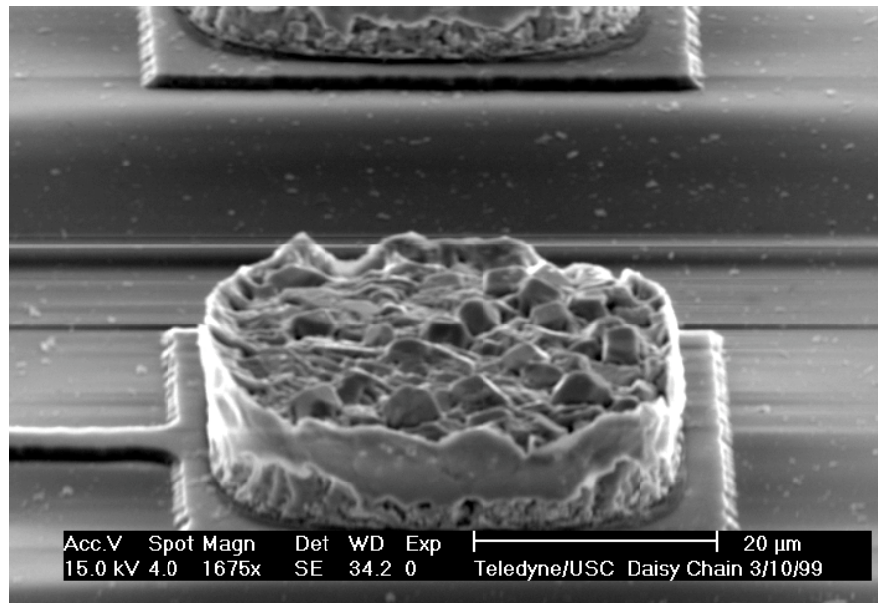
The first experiments that we undertook during the initial program period were designed to test the feasibility of employing single-sided bump contacts using thermally evaporated indium bumps instead of the more traditional dual-bump structure. This unusual approach is dictated by our desire to eventually be able to use commercially-available control, DSP, microprocessor, and DRAM chips in system-level implementations, as well as ASICs designed and fabricated by other vision groups worldwide. Often, such commercially-produced chips are available only as single die and not in wafer form, making indium bumping of each individual die an expensive and undesirable proposition. Although the indium bump deposition process is relatively benign, process incompatibilities can potentially limit the range of choices of both silicon chips and gallium arsenide chips that can be flip-chip bonded using the dual-bump (deposition on both substrates) structure.

During the initial stages of the research program, several four-inch silicon (Si) wafers containing a patterned flip-chip structure referred to as a "daisy chain" were conformally covered with indium bumps that were deposited in arrays designed to match the pre-patterned electrode arrays on the Si wafers. The daisy chain structure is composed of a 3 cm by 1 cm "base" or "bottom" chip and a smaller 1 cm by 1 cm "top" chip. A  $40 \times 40$  array of pairwise-



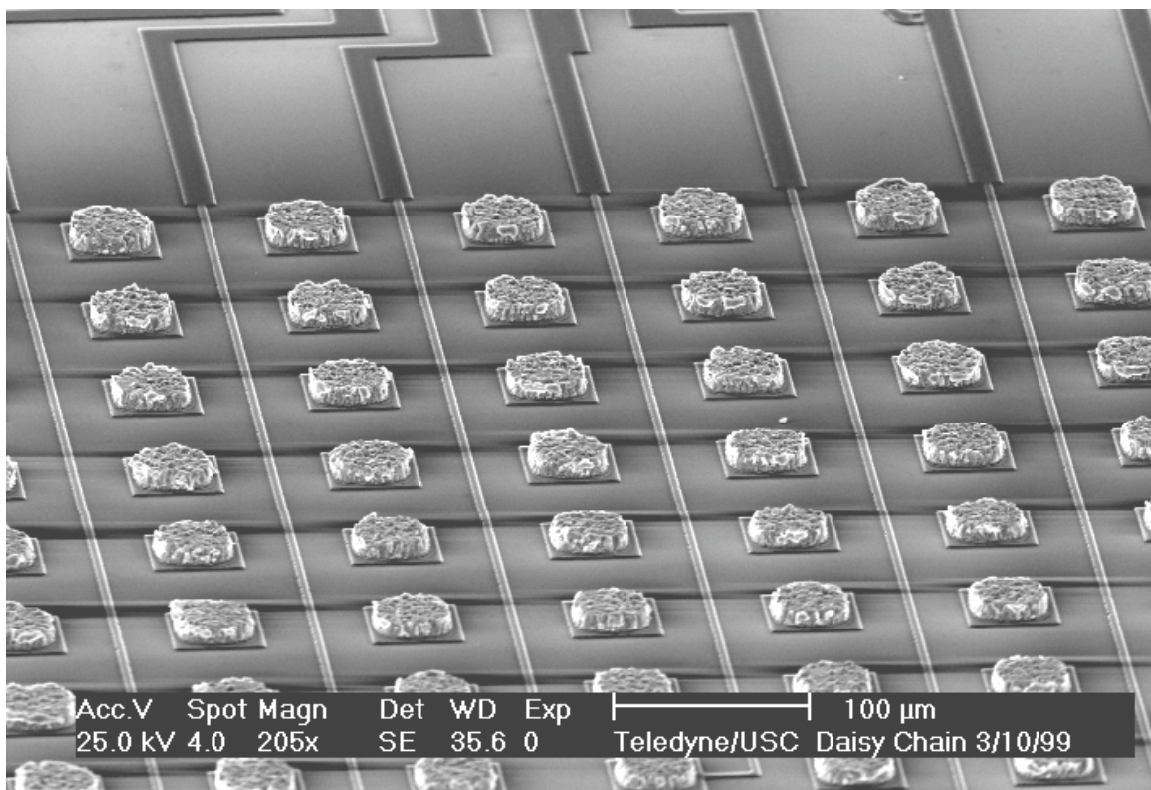
interconnected aluminum electrodes in the middle of the base chip is physically aligned with the top chip's corresponding  $40 \times 40$  array of pairwise-interconnected but otherwise isolated aluminum electrodes using the flip-chip bonder. After the top and bottom chips have been bonded together, a  $40 \times 40$  array of top-to-bottom electrodes results, interconnecting the two chips electrically and providing many test patterns that include from 2 to 40 indium bumps in each independently accessible pattern. Each individual pattern is accessible from the edge of the bottom chip, where large connection pads are provided for wire-bonding or probing. This configuration allows for a number of tests to be performed, ranging from basic electrical continuity (in multiple configurations) to measurement of the indium bump connection impedance as a function of frequency over the range of interest.

Two configurations using single-sided indium bumps were tested during the research program. The first experimental configuration incorporated indium bumps deposited on the base chip and bonded to an unbumped top chip, while the second experimental configuration was composed of an unbumped base chip bonded to a indium bumped top chip.



**Fig. 38.** Scanning electron micrograph of a single indium "velcro" bump deposited on an aluminum bonding pad, and interconnected to a wire-bonding pad by means of a metallization line. The surface morphology of the indium bumps is seen to be polycrystalline in nature, with large-scale RMS variations in bump height. In combination with the clearly visible sharp corners and edges on the bump surface, these large-scale variations promote low-impedance contacts through penetration of the native indium oxide layer.

Tests were successfully performed mating an unbumped top chip to an indium bumped base chip at bonding pressures of 10,000 grams (6.25 grams per bump) and 8,500 grams (5.3 grams per bump), with successful mechanical mating achieved in each case. Initial electrical testing indicated unusually high ( $k\Omega$ ) resistance values for small current/voltage signal levels flowing through the as-bumped device, with much lower impedances in the few tens of ohms range observed for voltages above 1 V. This is the result of either a native indium-oxide barrier on top of each indium bump, and/or an aluminum oxide layer on the aluminum bonding pads on both die substrates. Etch-before-bonding experiments were performed to lower the characteristic impedance of each interconnection bond. While partially successful, this approach does not represent an optimal solution.



**Fig. 39.** Scanning electron micrograph of one portion of a  $40 \times 40$  array of indium bumps deposited on aluminum bonding pads, and interconnected to wire-bonding pads by means of metallization lines. The array is designed to incorporate a daisy-chain pattern, allowing for direct measurement of bump contact resistance in varying-length sequences of indium bumps.



During the research program, we extended these initial flip chip bonding characterization experiments in two parallel directions: (1) continued development of appropriate processing sequences for indium bump contacts between aluminum bonding pads (important for the extension of this research to typical industrially-produced chip sets); and (2) the initial development of an appropriate processing sequence for indium bump contacts from either aluminum or gold pads (on the as-deposited side of the bump contact) to gold pads (on the flip-chip-bonding interface side of the bump contact).

In the first case, that of aluminum bonding pads, additional flip chip bonding experiments with chip-on-glass configurations that employ indium tin oxide patterned electrodes and electrode pads identified oxidized aluminum bonding pads as the source of the high resistance contacts observed at low voltages (as described above), and not a native indium oxide as might otherwise be suspected. We addressed this issue by investigating a zincation approach to pre-treat the aluminum contact pads, thereby avoiding the native aluminum oxide layer.

In the second case, gold bonding pads were chosen for the counterelectrodes in a number of continuing experiments. Although these are nontraditional bonding pads for most industrially-produced Si CMOS ASIC's, initial experiments showed dramatically lowered resistances at very low voltages, with excellent long term stability. During the research program, we continued to pursue this approach for flip chip bonding of Si CMOS driver chips to VCSEL arrays, as the VCSEL arrays can be fabricated using gold pads as a final deposition step.

### **Development of High Refractive Index Diffractive Optical Elements (DOEs)**

As both the silicon (Si CMOS driver circuit) and gallium arsenide (modulator or VCSEL array) layers are relatively high index, it may prove advantageous to fabricate the DOE arrays in gallium arsenide or silicon as well, on the basis of optimal (optical) impedance matching. A second advantage of this approach over the use of more traditional glass or quartz substrates is the necessity of efficient heat removal (especially in the case of the VCSEL arrays), in conjunction with the need for well-matched thermal coefficients of expansion to maintain alignment over the anticipated operating temperature range.

To this end, we undertook to design, fabricate, and test DOE arrays implemented in substrate materials with high indices of refraction. As a first test case, we used a combination of optical lithography and electron cyclotron resonance (ECR) etching to fabricate 4:2:1 fan-out patterns in GaAs substrates, as shown in Fig. 40. In the 4:2:1 fan-out pattern, a  $3 \times 3$  array of interconnections is created, placing four units of diffracted intensity into the zeroth order (corresponding zeroth nearest neighbor), two units of diffracted intensity into each of the four nearest neighbors, and finally placing a single unit of diffracted intensity (all in relative units) into each of the four next nearest neighbors. As the optimal diffraction efficiency is obtained when the optical path difference

between the etched and unetched regions is equal to a half-wave ( $\pi$  phase shift) at the design wavelength, the design etch depth was 1920 Å, and the measured etch depth was 1978 Å, as measured with a Sloan Dektak II surface profilometer. In this fabrication sequence, the ECR etching was performed in a PlasmaQuest Model 98 ECR etcher using both BCl<sub>3</sub> and Ar gas sources, a DC bias of 100 V, an RF power of 300 W, and with currents of 170 A and 80 A supplied to the upper and lower magnetic sources, respectively. An SEM photograph depicting the resulting side-wall profile is shown in Fig. 41.

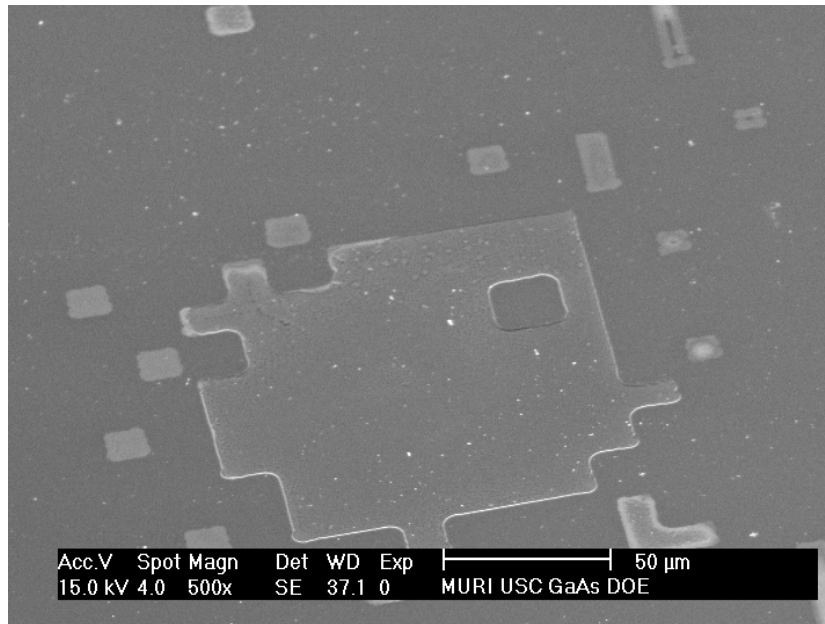
The diffraction efficiencies of each of the nine diffracted orders were measured at the design wavelength of 980 nm, resulting in errors from the theoretical (relative) diffraction efficiencies of between 1% and 24%, depending on the diffracted order, with an average error magnitude of 11.5% (treating all error deviations as positive quantities regardless of sign). Although these errors are likely marginally acceptable in a neural network environment, we decided to approach the reduction of these errors by antireflection coating both the back and front sides of the DOE elements in order to eliminate multiple internal reflections and their associated interference terms, as described in more detail below.

### **Development of High-Performance Antireflection Coatings for High Refractive Index DOEs**

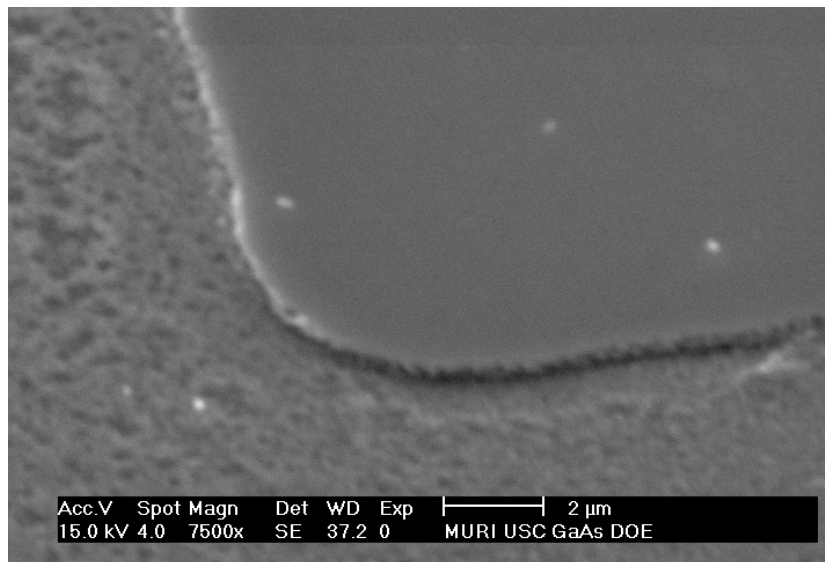
The hybrid electronic/photonic packaging scheme proposed for the implementation of an adaptive optoelectronic eye involves multiple vertically-interconnected layers of silicon VLSI detector/processing chips with interleaved layers of III-V compound semiconductor modulators or VCSELs, as well as layers of diffractive optical elements (DOEs) and possibly microlens arrays. In the modulator case, an optical power bus layer is also included.

During the research program, we undertook a study of the potential deleterious effects of multiple reflections on the integrity of the dense fan-out/fan-in optical interconnections in such a multilayer stack with up to eight interfaces between the sources and their corresponding detectors. The conclusion of this study was that multiple reflections can in fact pose a severe problem, requiring the antireflection (AR) coating of all of the layers in the multichip module. Several of these layers contain active photonic devices, and hence optimally require a combined AR coating and electrically conductive contact. In addition, at least one of the layers will contain a diffractive optical element that is characterized by a highly nonuniform surface (as described above).

We previously developed a transparent conductive coating for GaAs layers that can also provide a high-performance antireflection coating function as well [Karim, 1993]. During the research program, we evaluated the potential of this type of coating for application to the highly nonuniform surfaces characteristic of diffractive optical elements.



**Fig. 40.** Scanning electron microscope (SEM) photomicrograph of a GaAs diffractive optical element (DOE) sub-element within a  $20 \times 20$  array of identical sub-elements, fabricated by means of optical lithography and ECR etching techniques.



**Fig. 41.** An SEM photomicrograph of the GaAs DOE sub-element shown in Fig. 40, tilted in this case to show the uniformity and verticality of the as-ECR-etched side-wall profile.

During the research program, we designed, deposited, and evaluated the performance of such AR coatings on DOEs fabricated in both silicon and gallium arsenide. A 4:2:1 DOE fan-out pattern etched in a GaAs substrate as described above was coated on both front and back surfaces with a 1296 Å layer of indium tin oxide (ITO), deposited by RF magnetron sputtering using a Sloan S-310 Sputtergun at a pressure of 13 µm Hg, comprising 55 sccm of 99.9% Ar and 0.1% O<sub>2</sub>, at an RF power level of 250 W. As compared with the previous result (described above) on the uncoated GaAs DOE array with errors from the theoretical (relative) diffraction efficiencies of between 1% and 24%, depending on the diffracted order, with an average error magnitude of 11.5% (treating all error deviations as positive quantities regardless of sign), the antireflection (AR) coated GaAs DOE array exhibited errors from the theoretical (relative) diffraction efficiencies of between 0.6% and 9.5%, depending on the diffracted order, with an average error magnitude of 5.2%, a substantial improvement.

In this experiment, the first surface reflectivity of the native GaAs substrate (31.8%, due to an index of refraction 3.52 at 970 nm) was reduced to 0.4%. The resulting antireflection coatings are robust, broadband, and relatively easy to tune in order to match the index of refraction of the substrate over a range of design wavelengths (for a single layer coating, the optimal index of refraction is the square root of the substrate index of refraction). In addition, enough design flexibility is afforded to allow for the antireflection coating of other high index of refraction substrates (such as silicon).

### **Design and Fabrication of Low Threshold Vertical Cavity Surface Emitting Laser Arrays**

One of the most exciting applications of the vertical cavity surface emitting laser (VCSEL) is in free-space optical interconnections at the chip-to-chip level. Hybrid integration of optoelectronic devices onto Si-based systems is a promising solution for achieving higher performance in computer systems or dense optical interconnections. Bottom-emitting VCSEL's are particularly suitable for hybrid integration with silicon VLSI chips using a flip-chip bonding technology. A large two-dimensional VCSEL array could be transferred in one procedure, thereby reducing the device capacitance. Bottom-emitting VCSEL arrays are mainly operated at 980 nm to make use of the transparency of GaAs substrate at this wavelength. In this section of the report, the structure and the fabrication of 980 nm VCSEL arrays on GaAs substrate will be discussed. In a subsequent section, the result of hybrid integration with Si VLSI neuron unit arrays will be presented.

A key issue that limited the potential use of vertical cavity surface emitting laser arrays in the hybrid MCM structures described in this report is the high power dissipation associated with current-generation VCSEL's. This high power dissipation can be accommodated by either using large spatial separations of the VCSEL's (which reduces the area interconnection density), or by using a low duty cycle (which reduces the total number of connections per second that can be implemented). In order to reduce these undesirable effects, and to instead make use of the relative architectural and device simplicity offered by vertical cavity

lasers, it is important to develop low threshold current (low power dissipation) VCSEL's that can be integrated with silicon driver chips by means of flip-chip bonding techniques.

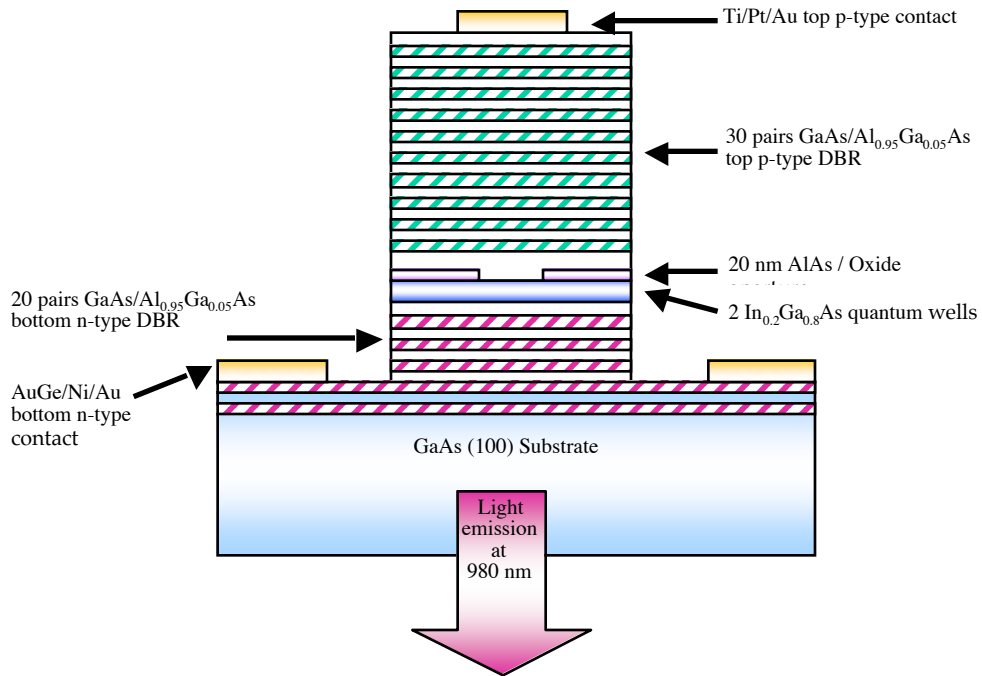
In a collaborative research effort on VCSEL's with Prof. P. Daniel Dapkus' group at USC, a better understanding of the role played by the placement of the aluminum oxide aperture in determining the scattering losses in VCSEL cavities was achieved during the research program. Use of optimized aluminum oxide apertures resulted in demonstrated threshold currents as low as 52 mA, and also led to designs in which the VCSEL characteristics are much less sensitive to variations in the oxidation length than previously existed.

### VCSEL Structure

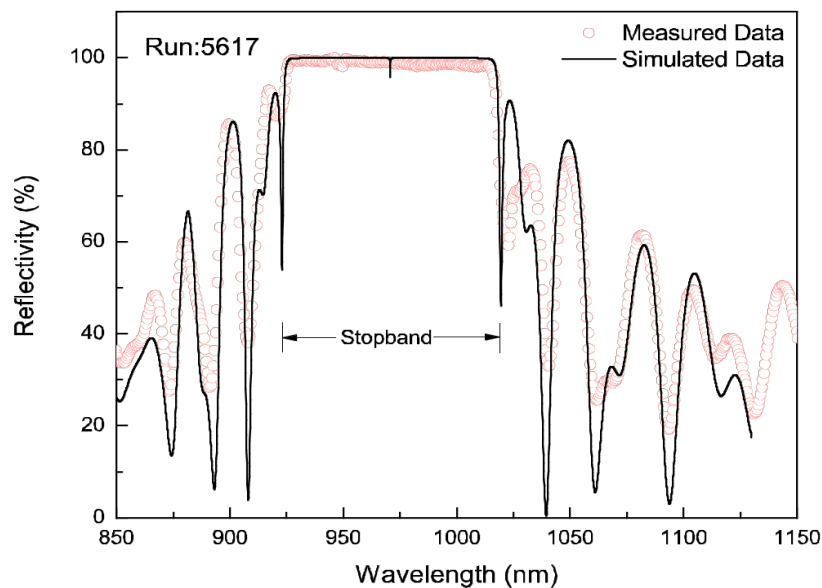
The typical components of a VCSEL are two high-reflectivity DBR mirrors that surround the optical cavity, and the gain medium used for light emission. Figure 42 shows a schematic cross section of a 980 nm VCSEL structure. The VCSEL structures were grown on (100) semi-insulating GaAs substrates using a low-pressure metal-organic chemical vapor deposition (MOCVD) reactor. The structure consists of: 20 pairs of *n*-type  $\lambda/4$   $\text{Al}_{0.95}\text{Ga}_{0.05}\text{As}$  and  $\lambda/4$  GaAs bottom DBR (99.4% reflectivity); two 7.8 nm thick  $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$  quantum wells; a 20 nm AlAs layer to be oxidized for the current aperture; 30 pairs of *p*-type  $\lambda/4$   $\text{Al}_{0.95}\text{Ga}_{0.05}\text{As}$  and  $\lambda/4$  GaAs top DBR (99.99% reflectivity); and a 10 nm thick *p*+ GaAs cap layer.

For both DBR mirrors, GaAs and  $\text{Al}_{0.95}\text{Ga}_{0.05}\text{As}$  are used as high and low index materials, respectively. Between them, a 20 nm thick  $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$  grading layer was used. They were uniformly doped with the concentration of  $1 \times 10^{18} \text{ cm}^{-3}$  for both *n*-type and *p*-type DBRs. Only the first five pairs of *p*-type DBR adjacent to the cavity were lightly doped ( $\sim 5 \times 10^{17} \text{ cm}^{-3}$ ) to reduce the free carrier absorption. Two 7.8 nm thick  $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$  quantum wells were used as a gain medium, and the quantum wells were separated by a 15 nm thick GaAs barrier to prevent dislocation formation. A 20 nm thick AlAs layer was located in the first period of the *p*-type DBR and aligned with the node of the standing wave pattern to minimize the scattering loss after it is oxidized.

Figure 43 shows the reflectivity spectrum taken from the sample after growth and the corresponding calculated data. The width of the mirror stopband is proportional to the refractive index difference between the high and low index materials in the DBR mirrors. If the difference in index of refraction between the DBR pairs is small, the stopband is narrower. The position of the stop band and the side peaks of the measured spectrum match the simulated one quite well. The measured PL spectrum peak from this wafer was 975 nm.



**Fig. 42.** A schematic cross-sectional diagram of the 980 nm bottom-emitting VCSEL structure with InGaAs quantum wells in the active region.



**Fig. 43.** VCSEL reflectivity spectrum showing the mirror stopband and the cavity resonance at 980 nm. The measured and simulated reflectivity of a completed 980 nm VCSEL is shown.

## VCSEL Fabrication

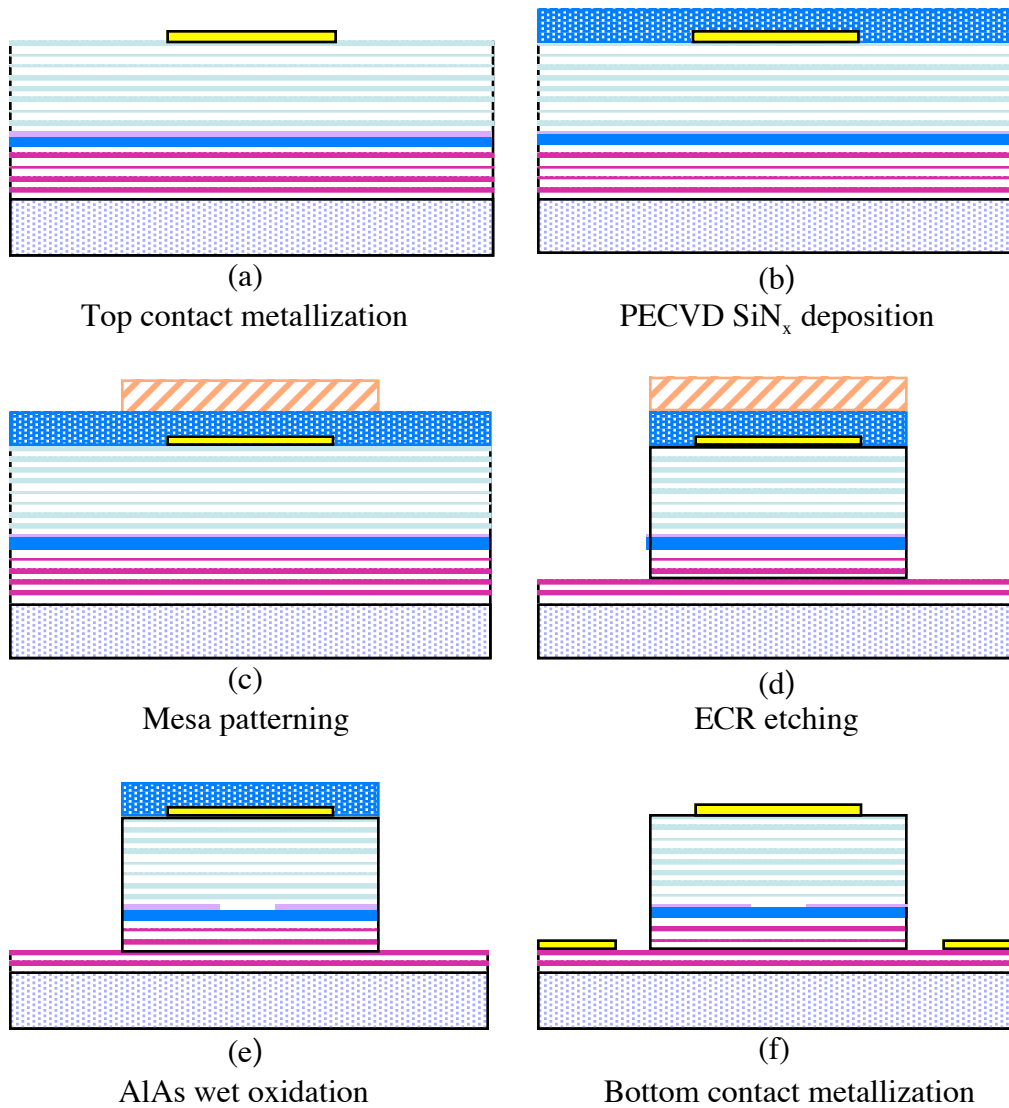
For the fabrication of 980 nm VCSEL arrays, three different mesa sizes of masks for VCSEL arrays were used; 20  $\mu\text{m}$ , 35  $\mu\text{m}$ , and 50  $\mu\text{m}$ . In a 20  $\times$  20 VCSEL array, each mesa was placed on a pitch of 125  $\mu\text{m}$ . Figure ZZ(a)-(f) shows schematically the 980 nm bottom-emitting VCSEL device processing procedure. The sample was first cleaned with TCE, acetone, methanol, and DI water to remove any contaminants on the surface. A 20  $\mu\text{m} \times 20 \mu\text{m}$  square metal lift-off mask was patterned with AZ5214 photoresist. An image reversal method was used to give an undercut edge profile so that the contact metal could be lifted off easily. After the pattern was developed, the sample was treated with oxygen plasma ashing to remove the photoresist residue on the developed surface. Before loading the patterned sample into the metal evaporator, the sample was cleaned with diluted hydrochloric acid,  $\text{HCl}:\text{H}_2\text{O} = 1:10$ , to remove any native oxide on the GaAs surface.

A 30 nm/50 nm/200 nm thick Ti/Pt/Au multilayer metal contact was deposited in an e-beam evaporator vacuum chamber to form *p*-type ohmic contacts to the GaAs surface. After the metallization process was completed, the sample was soaked in acetone for 5 minutes to lift off unwanted metal layers.

A 200 nm thick  $\text{SiN}_x$  thin film was deposited by PECVD at 275  $^\circ\text{C}$  to protect the top *p*-type metal during wet oxidation, and to be used as an etch mask during ECR etching. A 35  $\mu\text{m} \times 35 \mu\text{m}$  photoresist (AZ5214) square was patterned to define the mesa for subsequent  $\text{SiN}_x$  RIE etching. The  $\text{SiN}_x$  thin film was then etched in an RIE system using a  $\text{CF}_4$  plasma (100 mTorr, 100 W, 2 min) to create a 35  $\mu\text{m} \times 35 \mu\text{m}$  square mesa etch mask for a 20  $\times$  20 VCSEL mesa array during the following ECR etching process.

It is important to have a uniform mesa size across the array, as well as vertical sidewalls for the VCSEL mesas. If the mesa size or the vertical sidewalls are not uniform, then the wet oxidation of AlAs will not be uniform, either. Therefore, anisotropic dry etching for mesa patterning is preferred, since it offers improved uniformity of the mesa size and the vertical sidewalls of mesa as compared with wet etching.

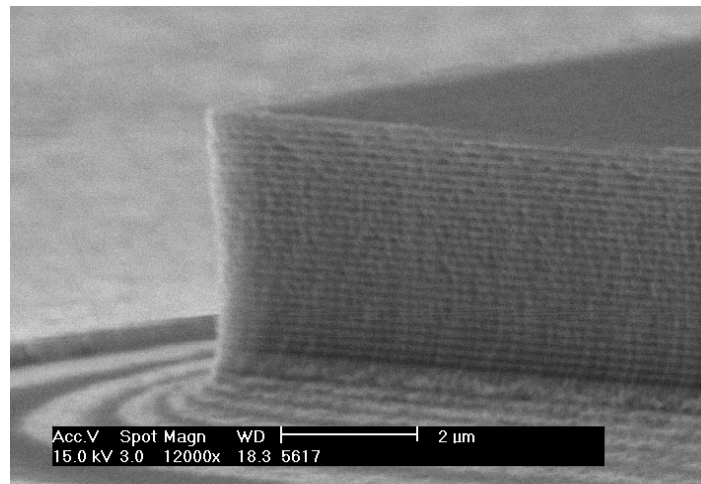
The VCSEL mesas were formed by plasma etching in an ECR dry etching system using  $\text{BCl}_3$  and Ar. Typically, VCSEL devices have very thick epitaxial structures, and require very accurate control of the etching depth. In order to etch VCSEL devices reproducibly, an *in-situ* laser reflectometry system was built to monitor the etching process. A He-Ne laser entered the ECR chamber through a quartz window at an angle of 60 $^\circ$  and reflected off the sample on the chuck through another quartz window to a silicon photodetector. The detected current was monitored using a Keithley picoammeter, which was connected to a chart recorder to produce an output record.



**Fig. 44.** Schematic diagram of the 980 nm bottom-emitting VCSEL array fabrication procedure. (a) Ti/Pt/Au top *p*-type contact metallization;  $20\ \mu\text{m} \times 20\ \mu\text{m}$  mask (b) PECVD  $\text{SiN}_x$  deposition (c) mesa patterning;  $35\ \mu\text{m} \times 35\ \mu\text{m}$  mask (d) ECR etching (e) AlAs wet oxidation (f) AuGe/Ni/Au bottom *n*-type contact metallization;  $50\ \mu\text{m} \times 50\ \mu\text{m}$  mask.



With this system, the VCSEL devices were etched at the desired target etch depth with a tolerance of less than  $\pm 2\%$ . The etching rate of GaAs/AlGaAs was  $\sim 0.30 \mu\text{m}/\text{min}$  with ECR etching conditions set to 300 W microwave power, 100 V RF dc bias, 30 sccm total flow rate ( $\text{BCl}_3/\text{Ar} = 25/5 \text{ sccm}$ ), upper/lower magnet = 170 A/40 A, and a 4 mTorr chamber pressure. During 22 minutes of ECR etching, 30 pairs of top *p*-type DBR, an AlAs layer, the quantum well region, and part of the bottom *n*-type DBR were etched for epitaxial wafer 5617. A total thickness of  $6.5 \mu\text{m}$  of epitaxial growth was etched. The ECR etching process was stopped at the tenth pair of the bottom *n*-type DBR below the cavity in order to expose the bottom DBR layer for broad bottom *n*-type contact metallization as well as to expose the edge of the AlAs layer for subsequent wet oxidation. Figure 45 shows an SEM image of the sidewall of a VCSEL mesa after ECR etching, showing that the etched mesa sidewalls were vertical and smooth.



**Fig. 45.** An SEM photomicrograph of the GaAs/AlGaAs DBR mesa structure after ECR etching (4 mTorr chamber pressure, 100 V RF dc bias,  $\text{BCl}_3/\text{Ar} = 25/5 \text{ sccm}$ , upper/lower magnet = 170 A/40 A, and 300 W).

After the edge of the AlAs current confinement layer was exposed by the ECR etching process, the sample was kept in methanol to avoid oxidation of the exposed AlAs layer by oxygen and moisture within the air. Usually, the sample was etched immediately before the wet oxidation process. The wet oxidation process was carried out in a 1 inch diameter open quartz furnace with 300 sccm  $\text{N}_2 + \text{H}_2\text{O}$  steam at  $425^\circ\text{C}$ . The  $\text{H}_2\text{O} + \text{N}_2$  environment was created by bubbling ultra high purity  $\text{N}_2$  gas through a 1 liter round bottom flask, which was maintained at a constant temperature of  $87.7^\circ\text{C}$ . The  $\text{N}_2$  flow rate was precisely controlled by an electronic mass flow controller. The gas lines between the water bubbler and the oxidation furnace were heated to a temperature of  $120^\circ\text{C}$  to prevent water condensation that could cause unstable oxidation rates.

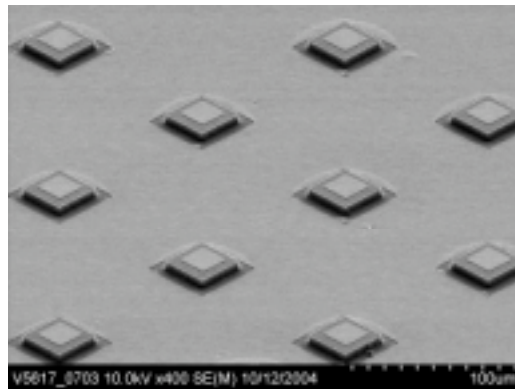
The samples were placed on a piece of Si in a quartz boat and then slid into the oxidation furnace with a quartz rod. All of the parameters (water bubbler

temperature, furnace temperature, and  $N_2$  flow rate) were carefully kept under the same conditions to obtain reproducible oxidation. Since the sidewalls of the mesas were exposed after the ECR etching, the oxidation fronts proceed laterally into the mesa center from the edges of the mesas. The oxide extent can be observed with an optical microscope if the layer structure is thin enough, while an IR camera is necessary for examining the oxide layer in a thick VCSEL structure. Since the bottom-emitting epitaxial structure was too thick to allow observation of the oxide extent with an optical microscope, a test oxidation sample was fabricated in parallel with the main sample.

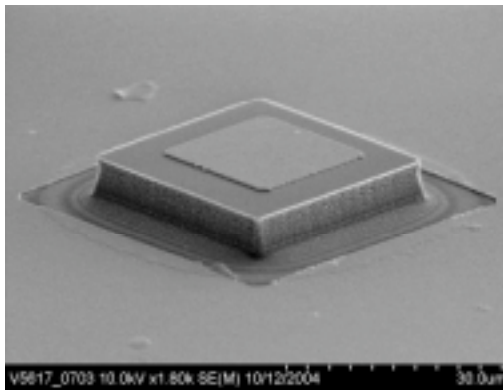
The oxidation rate was estimated from the test oxidation sample, and then applied to the oxidation of the main sample. The oxidation rate of the thin AlAs layer (20 nm) was  $\approx 0.62$  to  $0.68 \mu\text{m}/\text{min}$  under these conditions. Since the oxidation rate of the thin AlAs layer is slower than that of the thick AlAs layer, the thin AlAs layer was chosen for precise control of the oxidation rate. With these oxidation rates, we can control the final oxide aperture size in a VCSEL mesa structure to within  $1 \mu\text{m}$ .

Instead of immediately removing the sample from the oxidation furnace after the oxidation process was complete, the sample was moved to a lower temperature zone for 10 minutes. This two-step oxidation approach was employed to remove the intermediate by-products and to achieve stability of the resulting oxide.

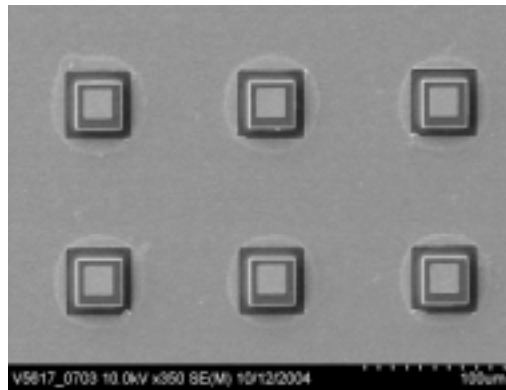
The  $\text{SiN}_x$  thin film on top of the VCSEL mesa was removed by using a  $\text{CF}_4$  plasma in an RIE etching system. Then a broad area AuGe/Ni/Au 100 nm/30 nm/100 nm  $n$ -type ohmic contact was evaporated onto the AlGaAs bottom  $n$ -type DBR layer to serve as a common anode. To alloy the contacts, the sample was loaded into a rapid thermal annealer (RTA), and then annealed at  $400^\circ\text{C}$  for 30 sec in forming gas with a  $1^\circ\text{C}/\text{sec}$  ramp rate. For bottom-emitting VCSEL's, it is necessary to apply an antireflection (AR) coating at the air and substrate interfaces to eliminate unwanted reflections. For this purpose, a  $\text{SiO}_x$  ( $n = 1.90$ ) layer was deposited in a Sloan e-beam evaporator for the 980 nm lasers fabricated on GaAs substrates. Without AR coatings, the L-I characteristics of the bottom-emitting VCSEL's showed strong feedback from the substrate-air interface. An SEM image of the completed device is shown in Fig. 46.



(a)



(b)

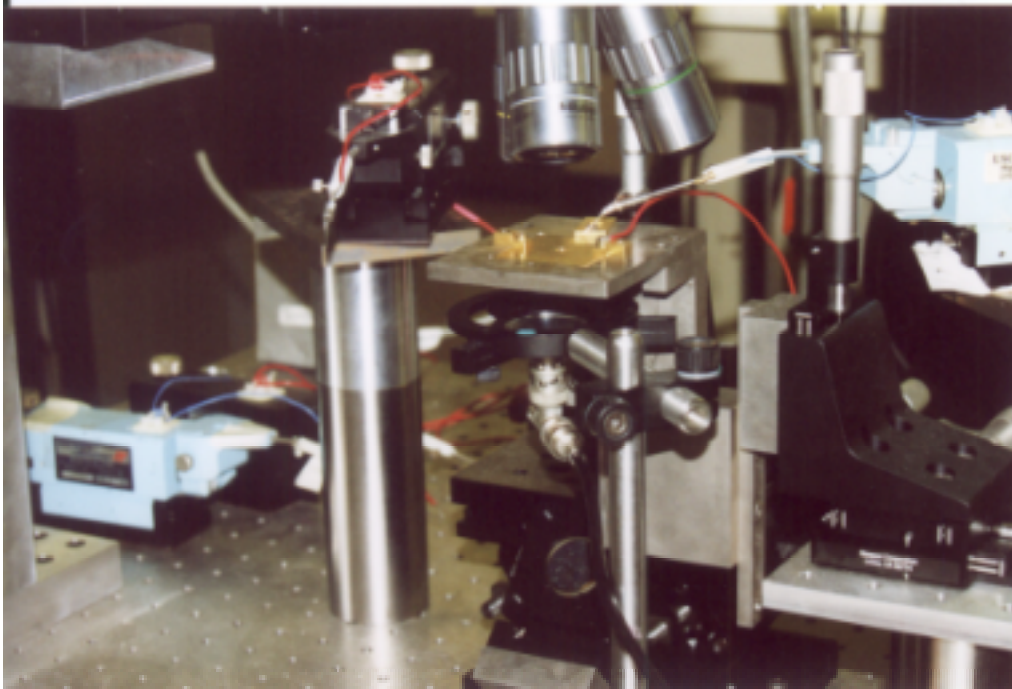


(c)

**Fig. 46.** SEM micrographs of a completed bottom-emitting 980 nm VCSEL array. (a) Side view of VCSEL mesas; each mesa is separated by a pitch of 125  $\mu\text{m}$ ; (b) Side view of a 35  $\mu\text{m}$   $\times$  35  $\mu\text{m}$  VCSEL mesa; (c) Top view of a VCSEL array.

### Bottom-Emitting 980 nm VCSEL Array Results

The completed VCSEL devices were tested at room temperature on an uncooled stage with continuous wave (CW) excitation. The CW measurements were made with an HP4142B modular DC source test instrument and a UDT calibrated broad area Si photodetector. Current-voltage (I-V) and current-light (I-L) measurements were made on individual devices. In addition, output emission spectra were recorded. The sample and the sample holder were located above a Si photodetector. Since these VCSEL devices were bottom-emitting, a 5 mm size hole was made at the center of the Au-plated sample holder. The majority of the fabricated VCSEL devices on the chip lie above open hole, which allowed the emission of light through the apparatus. The test setup for bottom-emitting 980 nm VCSEL's is shown in Fig. 47.



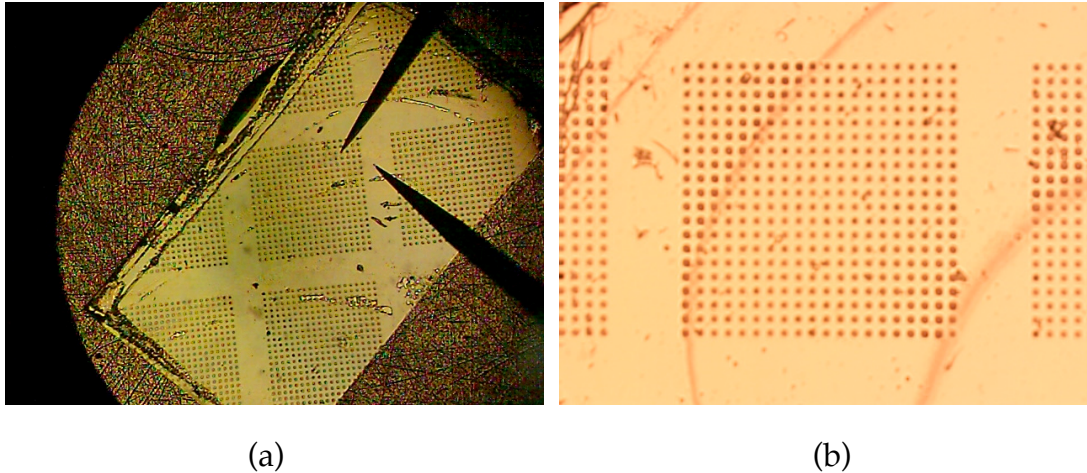
**Fig. 47.** The test setup for bottom-emitting 980 nm VCSEL's.

As may be seen in Fig. 48, one probe contacted the broad bottom metal layer, and the other probe was placed directly on top of a VCSEL mesa. Terra Universal tungsten probes were used for the VCSEL measurements. These probe tips were 1 inch long with a 0.0001 inch tip radius, a 5° taper angle, and an 0.23 inch taper length. These ultrathin probes allowed direct contact to VCSEL mesas of dimensions as small as  $20\ \mu\text{m} \times 20\ \mu\text{m}$ .

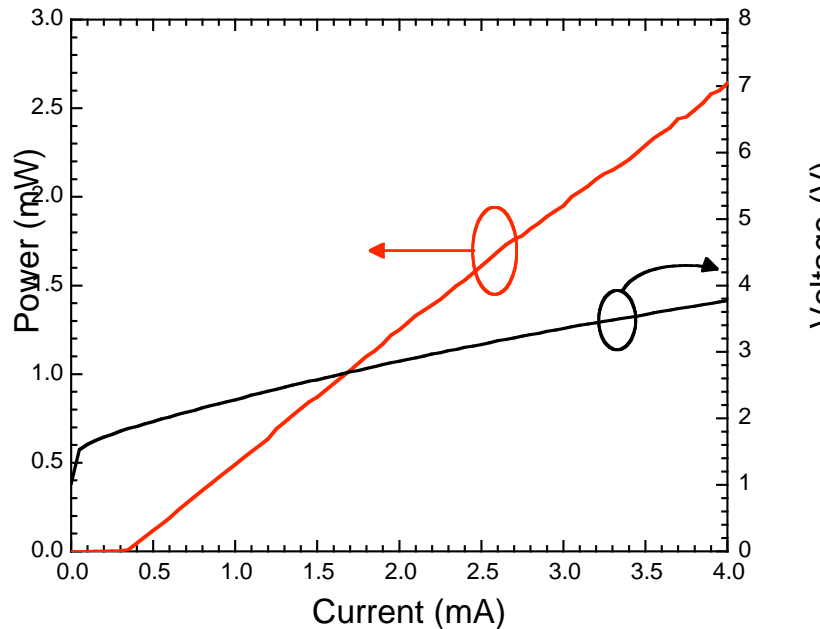
To measure the emission spectra and far field pattern, additional components were added to the measurement setup on an optical table. The light output from the laser passed through an opening in the sample holder and was collected by a 50X objective lens. The output was reflected by a mirror, and collimated by a simple lens. The optical signal was collected by a multimode fiber and analyzed by an optical spectrum analyzer. Figure 49 shows the typical L-I-V characteristics of a single VCSEL device within a  $20 \times 20$  VCSEL array. The threshold current and external quantum efficiency were  $350\ \mu\text{A}$  and 57%, respectively.

From Fig. 49, it can be seen that a VCSEL with a square oxide aperture  $\sim 3.5\ \mu\text{m}$  on a side has a threshold current as low as  $350\ \mu\text{A}$  with a threshold voltage of 1.6 V (2.0 V at the threshold current). The external quantum efficiency reaches 57% and the maximum power exceeds 3 mW at 5 mA. High efficiency and low threshold current are two desirable characteristics realize for large and densely packed VCSEL arrays. However, these two properties cannot be achieved at the same time, and tradeoffs between the wall-plug efficiency and the threshold current must be made when designing the VCSEL DBR mirrors.

For the hybrid integration of VCSEL arrays with Si circuitry, high wall-plug efficiency is more desirable than a low threshold current with a correspondingly low output power.



**Fig. 48.** CCD images taken through a microscope to show the probe geometry for 980 nm bottom-emitting  $20 \times 20$  VCSEL arrays; (a) top view of  $20 \times 20$  VCSEL arrays; (b) a higher magnification view of a  $20 \times 20$  array.

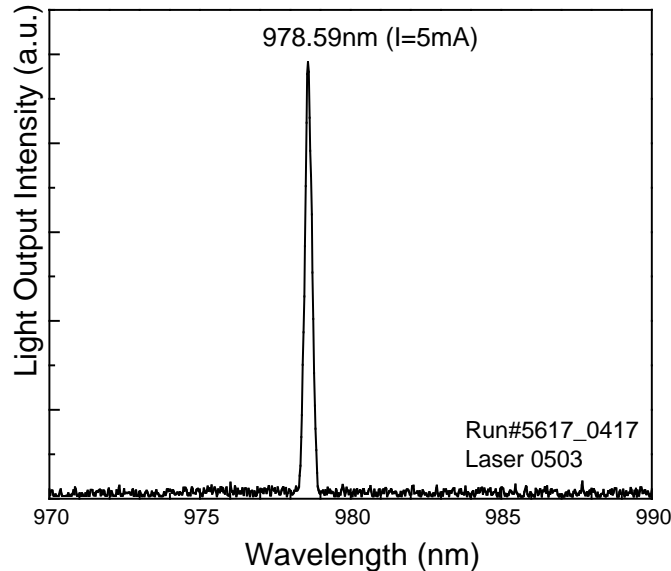


**Fig. 49.** The L-I-V characteristics of a 980 nm VCSEL from a  $20 \times 20$  VCSEL array under CW conditions. The threshold current and external quantum efficiency were  $350 \mu\text{A}$  and 57%, respectively.

A typical emission spectrum of a 980 nm bottom-emitting VCSEL array element is shown in Fig. 50. The VCSEL operated in single mode up to 5 mA



under CW conditions. Most of the VCSEL devices operated in single mode up to 10 times the threshold current ( $\sim 3$  to  $4$  mA). For this bottom-emitting laser, the center lasing wavelength was  $976.9$  nm at a current of  $1$  mA.



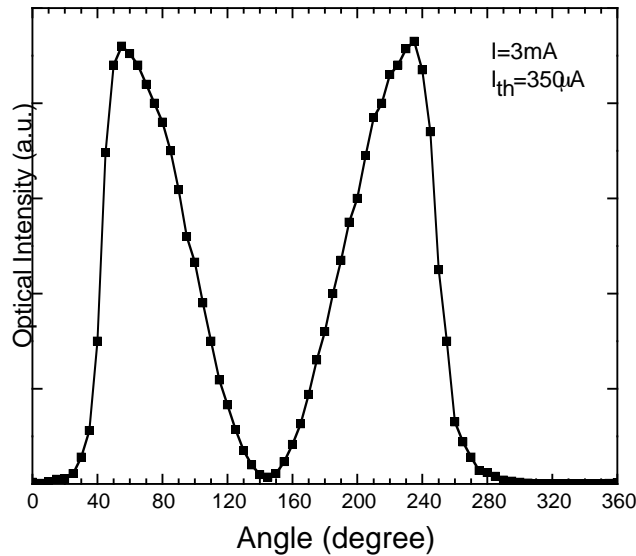
**Fig. 50.** Typical emission spectrum of a bottom-emitting VCSEL array element. The VCSEL device operates in single mode up to  $5$  mA under CW conditions.

Polarization measurements for the VCSEL arrays were performed by placing a  $980$  nm optical polarizer in front of the multimode fiber that coupled the optical signal into the optical spectrum analyzer. The measurement result is shown in Fig. 51. This data was taken at  $3$  mA current under CW conditions. From Fig. 51, it is clear that the VCSEL is linearly polarized.

The wavelength spectrum of a semiconductor laser is an important device characteristic, because in many applications spectral control of the laser output is required. Since the emission wavelength is determined by the reflectivity resonance, it is expected that the wavelength will shift due to, for example, ambient or internal temperature changes that occur in response to thermal change of the indices of refraction of the mirror. In addition to the wavelength shift induced by ambient temperature changes, we must also consider shifts induced by self-heating effects.

The output wavelength as a function of input current of one of the  $980$  nm VCSEL's is plotted in Fig. 52. The threshold current of this device is  $350$   $\mu$ A. We observe a wavelength shift with a rate of  $\sim 0.336$  nm/mA. The lasing wavelength shifts to longer wavelengths as the input current is increased. The observed wavelength shift is approximately linear in the power dissipated by the VCSEL. When VCSEL devices are in operation, they heat up due to the flow of current through the DBR mirror layers. As a result, the cavity resonance mode shifts to longer wavelengths due to changes in the refractive indices, and the relative positions of the cavity resonance and the gain spectrum will determine

the laser output wavelength for a given level of carrier injection. These results are very important for the development of matching DOE fanout patterns and PMCM device sizes that are tolerant to wavelength shifts of this magnitude, since the VCSEL's are driven with time-varying analog currents in the PMCM architecture.



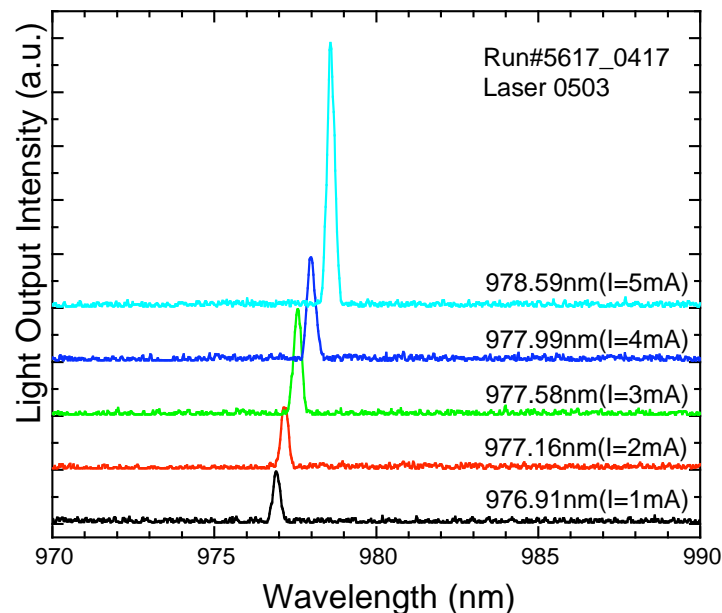
**Fig. 51.** The measured polarization properties of a 980 nm bottom-emitting laser. This data was taken at 3 mA under CW conditions with a 980 nm optical polarizer.

The L-I-V characteristics of an  $8 \times 8$  bottom-emitting 980 nm VCSEL array (125  $\mu\text{m}$  pitch) are shown in Fig. 53(a). The average threshold current is 321  $\mu\text{A}$  with a standard deviation of 23.2  $\mu\text{A}$ , and the average external quantum efficiency is 55.8% with a standard deviation of 1.96%. The maximum wall-plug efficiency approaches 23% at 1 mW output power, which is limited by the series resistance of  $\sim 460$  to 500  $\Omega$ . The maximum single mode optical output power is more than 2 mW under CW conditions.

In addition to the data shown in Fig. 53(a) from an  $8 \times 8$  array, key characteristics of a  $20 \times 20$  array of VCSEL's were also measured. Of the 400 lasers in the array, four lasers failed to lase and the contact pads of three lasers were destroyed before or during measurement. Figures 53(b) and (c) show histograms of the measured threshold current and external quantum efficiency of the  $20 \times 20$  VCSEL array. The uniformity of the  $20 \times 20$  array is not as good as that of the  $8 \times 8$  array. The standard deviations of the threshold current and efficiency are increased to 126.6  $\mu\text{A}$  and 3.7%, respectively. This non-uniformity was partly caused by the testing itself, as the L-I characteristics change in response to the amount of force applied to the device by the testing probe. Since

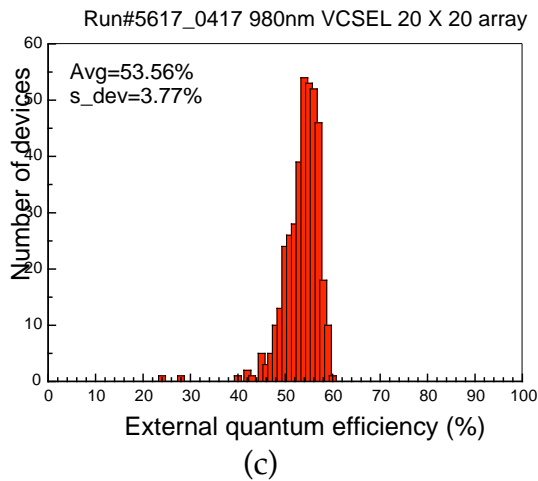
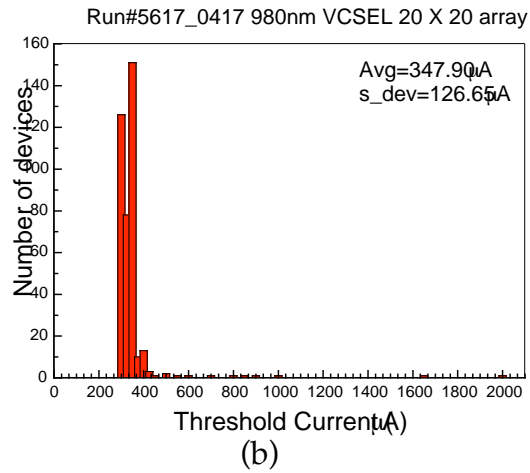
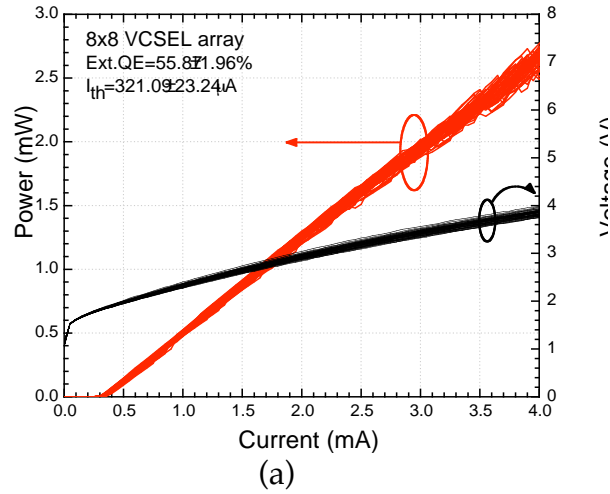
there was no extended contact pad for the VCSEL array top contact, the testing probe was directly placed on top of the VCSEL mesa during the device measurement. As a result, some of the VCSEL mesas were destroyed during the measurement due to the vibration of the setup. Also, the wet oxidation step caused non-uniformity of the oxide apertures in the  $20 \times 20$  VCSEL arrays during the fabrication process. The oxide apertures provide strong confinement for the optical modes in VCSEL structures, and the device performance varies with the absolute size of oxide aperture.

Control of the aperture size across the sample becomes very critical for the uniformity of the L-I-V characteristics in VCSEL arrays. The wet oxidation process starts from the edge of the VCSEL mesas and advances toward the mesa centers. Since the oxide aperture is not defined by photolithography, control of the wet oxidation process becomes very important for the performance of these devices. The oxidation rate of the AlAs layer strongly depends on the layer thickness, composition, and temperature distribution in the oxidation furnace. The uniformity of both VCSEL growth and wet oxidation processes needs to be accurately controlled in order to improve the uniformity of oxide-confined VCSEL's.



**Fig. 52.** Light output intensity as a function of wavelength from a 980 nm bottom-emitting laser from a  $20 \times 20$  VCSEL array with different currents, measured under CW conditions.





**Fig. 53.** Device characteristics of (a)  $8 \times 8$  and (b, c)  $20 \times 20$  980 nm bottom-emitting VCSEL arrays. (a) L-I-V curves of 64 VCSEL's; (b) histogram of the threshold current; (c) histogram of the external quantum efficiency.

## Photonic Multichip Module (PMCM) Integration

During the research program, an important experiment was conducted to demonstrate a key stage in the integration of the photonic multichip module (PMCM). In this experiment, two 970-nm wavelength, top-emitting vertical cavity surface emitting lasers within an  $8 \times 8$  array fabricated at USC were wire-bonded to a submount, which was then connected to two independent VCSEL device drivers. The lasers exhibited 400  $\mu\text{A}$  thresholds, and were addressed by the device drivers to operate either one at a time, or both simultaneously.

A quartz-substrate diffractive optical element ( $n = 1.457$ , fabricated by e-beam lithography at QPS, Inc.) that implements a 4:2:1 fan-out pattern as described previously was mounted in proximity to the VCSEL array, such that the output beams from both lasers intercepted the same DOE. Each laser was turned on independently, resulting in the desired  $3 \times 3$  fan-out pattern as shown in Fig. 54. Both lasers were then turned on to matched output intensities, and a lens following the DOE array was placed such that the output patterns from the two lasers overlapped, as shown in Fig. 55.

This output pattern demonstrates the key function of fan-in, and is the first such demonstration of its kind, to the best of our knowledge. The measured optical reconstruction pattern is shown in Fig. 56, from which error percentages could be estimated based on incoherent summation rules. The minimum error observed in the 14 resulting diffracted spots was 0.2%, and the maximum error was 16%, with an average error of 7% (treating all error deviations as positive quantities regardless of sign).

The same 4:2:1 fan-out pattern was also photolithographically defined in a GaAs substrate, as described previously, and this DOE element was also used to perform the same fan-out/fan-in experiment described above. Successful fan-in was achieved, and quantitative performance comparisons against the quartz-substrate DOE will be included in forthcoming publications.

Also during the research program period, we demonstrated additional fan-in from multiple laser sources, and measured the degree of error over a range of relative laser output intensities. Continuing these experiments will allow us to determine the possible complications that may result from simultaneous VCSEL operation within an array by either electrical or thermal cross talk.

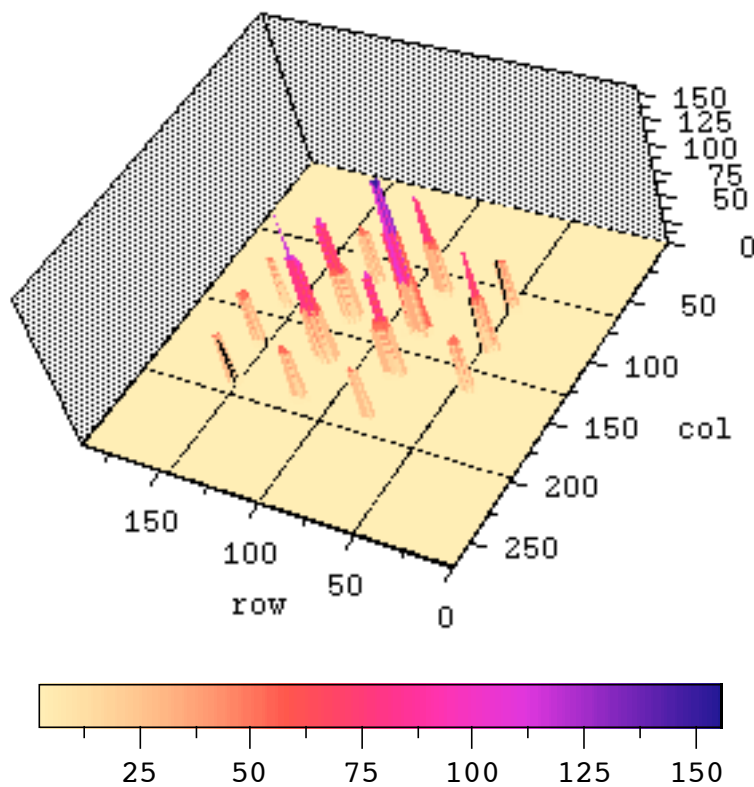
Vertical-cavity surface-emitting lasers (VCSELs) are an ideal light source for free-space optical processing since they can easily be fabricated into two dimensional arrays of individually addressed lasers for which the output beam is circularly symmetric with controllable divergence angle. However, it is a challenging process to integrate a large array of VCSELs onto a CMOS chip, since the electrical interconnections between the VCSEL and the CMOS electronic driver circuit must be short in order to reduce the parasitic capacitance, electrical cross talk, and the complexity of interconnect wires that would be required if a separate VCSEL was located some distance from the CMOS chip. As discussed in a previous section, the flip-chip bonding technique has previously been



**Fig. 54.** Fan-out pattern from a single 970 nm VCSEL element within an  $8 \times 8$  array, transmitted through a 4:2:1 diffractive optical element (DOE).



**Fig. 55.** Fan-in pattern from two 970 nm VCSEL elements within an  $8 \times 8$  array, transmitted through the same 4:2:1 diffractive optical element (DOE).



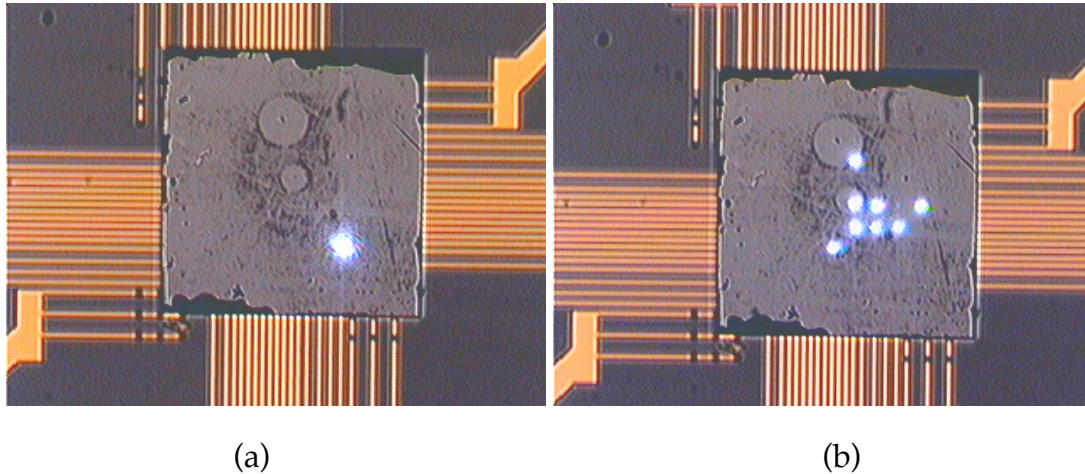
**Fig. 56.** Measured optical reconstruction pattern from two 970 nm VCSEL elements within an  $8 \times 8$  array, transmitted through the same 4:2:1 diffractive optical element (DOE). The spot size (FWHM) was 125 to 250  $\mu\text{m}$ , with a spot separation of 2.5 mm.

employed to integrate large arrays of GaAs multiple quantum well (MQW) modulator-arrays to CMOS circuits [Goossen, 1995]. By using a relatively simple flip-chip bonding technique in a similar manner, hybrid integration of bottom-emitting VCSELs array to silicon VLSI chips can be demonstrated. The integration of VCSEL arrays with gigabit-per-second CMOS circuits via flip-chip bonding technique has been demonstrated [Krishnamoorthy, 1999].

The co-integration of a VCSEL array with a Si drive chip (in this case a test chip) is illustrated in Fig. 57. As shown in the figure, two-dimensional arrays of VCSELs are flip-chip bonded on a pixel-by-pixel basis to the silicon VLSI chips, which act as VCSEL drivers through pinned out leads terminating in an array of bonding pads that are designed to match the VCSEL pitch.

The electrical and optical characteristics of the flip-chip bonded  $8 \times 8$  VCSEL array were measured, with results as shown in Fig. 58. The red curves (composed of square symbols) depict the pre-bonded characteristics of the laser, while the blue (smooth) curves depict the post-bonded characteristics of the laser. The output optical intensity as a function of drive current did not change

significantly, but the voltage-current characteristic changed slightly after the bonding process. The resistance of the bottom-emitting laser before flip-chip bonding was  $460\ \Omega$ . After bonding with the silicon mating substrate, the resistance of the laser was increased to  $540\ \Omega$ . These total electrical resistance values of the laser were determined by calculating the difference in slopes of the measured V-I curves before and after bonding.



**Fig. 57.** Optical micrographs showing operating flip-chip bonded VCSEL chip to a silicon mating substrate; (a) one laser in operation at  $I = 2\text{ mA}$ ; (b) 8 lasers operating simultaneously.

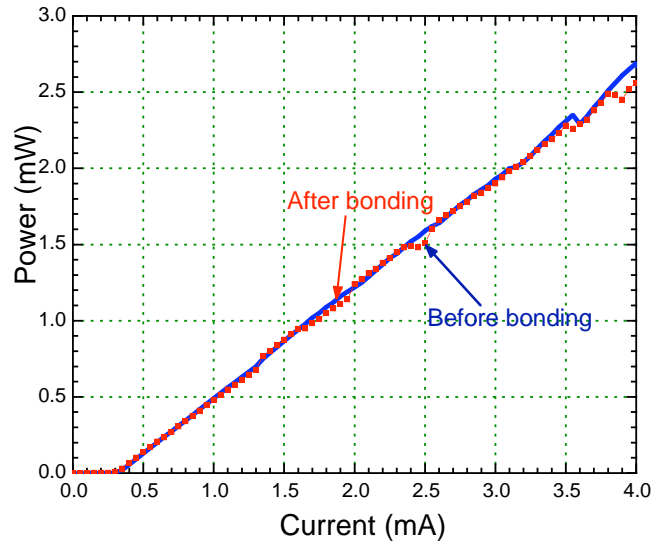
### Implementation of Variable-Kernel-Size Sobel Transformations

While pursuing a novel approach for the implementation of an adaptive optoelectronic eye, we remained committed to tracking the state-of-the-art in all-electronic implementations of related smart camera functions, and to making direct comparisons to projections for our emerging photonic multichip module.

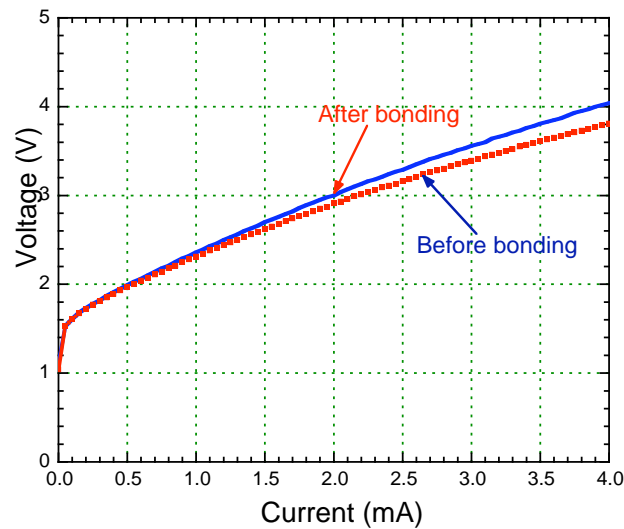
In order to achieve this goal, we undertook a study of the computational burden imposed by convolutional and nonlinear operations that are typical of image-processing or vision-related algorithms, as implemented on both emerging smart cameras and on desktop computers, from PC-scale through large-scale sophisticated workstations.

In these studies, we tested the implementation of a Sobel operation, commonly used in edge detection algorithms, on a variety of platforms. The Sobel operation can be cast in terms of two convolution operations and a (nonlinear) magnitude operation, and hence combines two key features of the emerging adaptive optoelectronic eye architecture. This operation was tested using  $3 \times 3$  convolution kernels across a standard  $256 \times 256$  pixel test gray-scale image. Results ranged from 1.181 seconds on an HP (Apollo) Series 700

workstation, through 0.109 seconds on an SGI R10000 workstation, to 0.030 seconds on an IVP MAPP 2200 smart camera (manufactured by IVP, Sweden) that incorporates 256 single pixel processing elements onboard the CCD imaging chip.



(a)



(b)

**Fig. 58.** Measured L-I-V characteristics of the flip-chip bonded VCSEL (a) light-current characteristic of the bonded VCSEL; (b) voltage-current characteristic of the bonded VCSEL (pre-bonded result = red (dotted) curves; post-bonded result = blue (smooth) curves).

Although the preliminary IVP MAPP 2200 smart camera data looked promising, several key disadvantages of this approach are noteworthy. The number of programmable filter operations for the current version of the smart camera was limited, there was no color support, only 127 bits of RAM were

available for each processing element, there was no support for conditional logic (if/then) statements, there was no way to program individual processors, and finally the smart camera proved difficult to program, as assembly language was required. Furthermore, the achievement of real-time frame rates with a Sobel operation is impressive, but the scaling to larger kernel sizes is not. In fact, kernel sizes larger than  $3 \times 3$  were not supported, and the range of achievable filter operations was predetermined by the limited number of kernels hardwired into the camera head.

### **PMCM Optical Power Budget**

In order to estimate the overall power requirements of the PMCM hardware implementation, a detailed analysis of the overall PMCM optical power budget was performed.

The VCSEL-based PMCM stack containing two adjacent VCSEL's illuminating a single DOE substrate in a simple two-layer PMCM device is shown schematically in cross section below. The output optical beam width at each substrate is initially determined by the VCSEL oxide confinement size, and later by the substrates' refraction properties as predicted by Snell's Law. This figure also shows a schematic representation of the total optical power as one progresses through the PMCM stack.

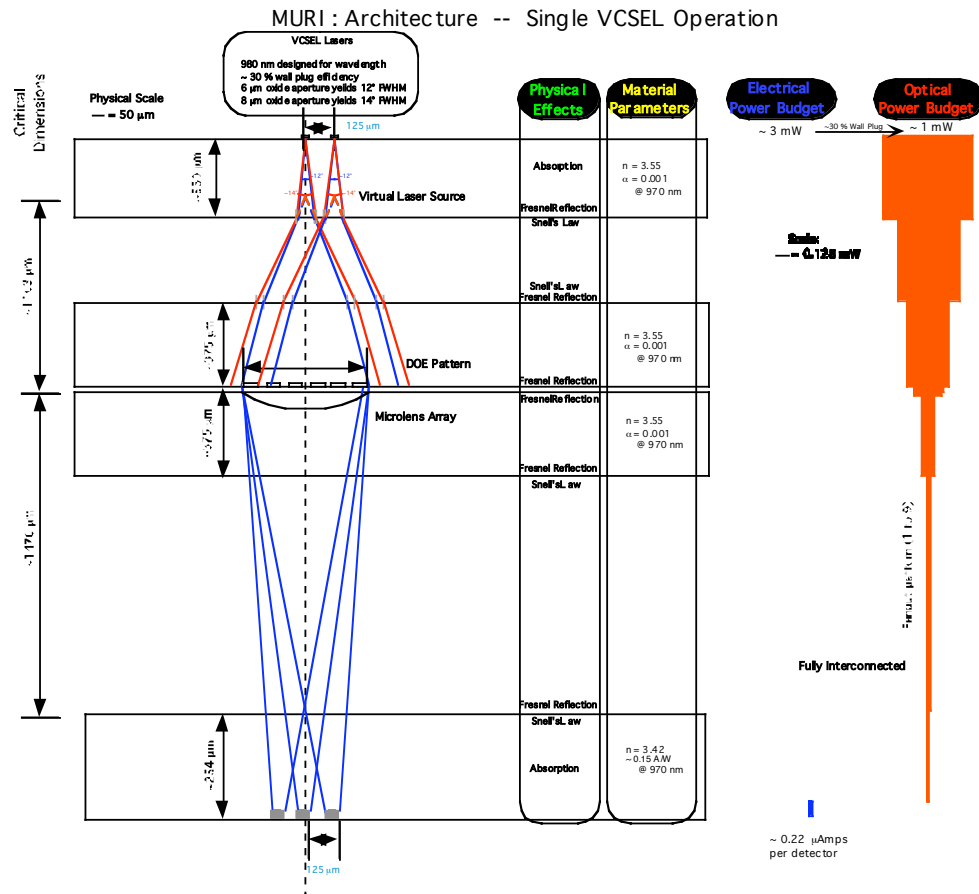
As configured for this calculation, each VCSEL is assumed to have a wall plug lasing efficiency of 33%, and a maximum output optical power of 1 mW (chosen due to the availability of VCSEL's with these parameters). This model also takes into account all Fresnel reflection and absorption properties of the substrates, while ignoring all unwanted optical diffraction effects. Furthermore, this model assumes that all light is normal to each surface, excludes any dispersion-based effects, and does not include multiple reflections within the stack architecture. With these effects notwithstanding, this model still provides an excellent starting point for initial optical power estimates needed for operation of the PMCM.

The two VCSEL's are separated by a  $125\text{ }\mu\text{m}$  pitch, with the red-lines representing the optical beam paths of an  $8\text{ }\mu\text{m}$  VCSEL oxide aperture, and the blue lines representing the optical beam path for a  $6\text{ }\mu\text{m}$  VCSEL oxide aperture. Both beams propagate through the GaAs VCSEL substrate with little loss until striking the back VCSEL surface. At this point, two optical effects occur. First, a Fresnel loss decreases the total optical power by nearly 33%. The second optical effect is the refraction properties caused by the difference in optical indices as characterized by Snell's Law. In our case, as the optical ray leaves a high index substrate and enters a lower index substrate (air), the optical beam is refracted away from the interface surface normal.

As shown in the figure, the optical paths for both adjacent VCSEL's exit the GaAs VCSEL substrate and enter the shared DOE substrate. The DOE substrate consists of a  $375\text{ }\mu\text{m}$  thick GaAs material in this configuration and is  $1103\text{ }\mu\text{m}$  away from the VCSEL substrate. The distance between these two substrates is



chosen to correspond to the silicon detector dimensions (25  $\mu\text{m}$  square for this exercise) as predicted by Fourier optics principles. Generally, the greater the illumination area of the incident optical beam, the smaller the (FWHM) diffracted optical beam profile. In this figure, the same area of the DOE pattern is simultaneously illuminated by the two adjacent VCSEL's.



**Fig. 59.** Schematic diagram of the PMCM characterizing all of the critical optical properties. The width of the orange line represents the total optical power at that point in the PMCM stack.

After passing through the DOE, a secondary substrate configured as a lens array is needed to perform the necessary Fourier transform. In previous figures in this report, the DOE and the lens array were integrated into the same substrate. This was achieved by creating a refractive index distribution that can simultaneously function as a lens, within which the DOE can be contained. For convenience, the optical power budget model breaks these two necessary functions into separate substrates. The lens array in this model is a surface relief pattern on a GaAs substrate. As configured, a single lens is used to perform the Fourier transformation of the DOE. Additional effects caused by the adjacent VCSEL light output including lens aberrations, non-normal Fresnel reflections,

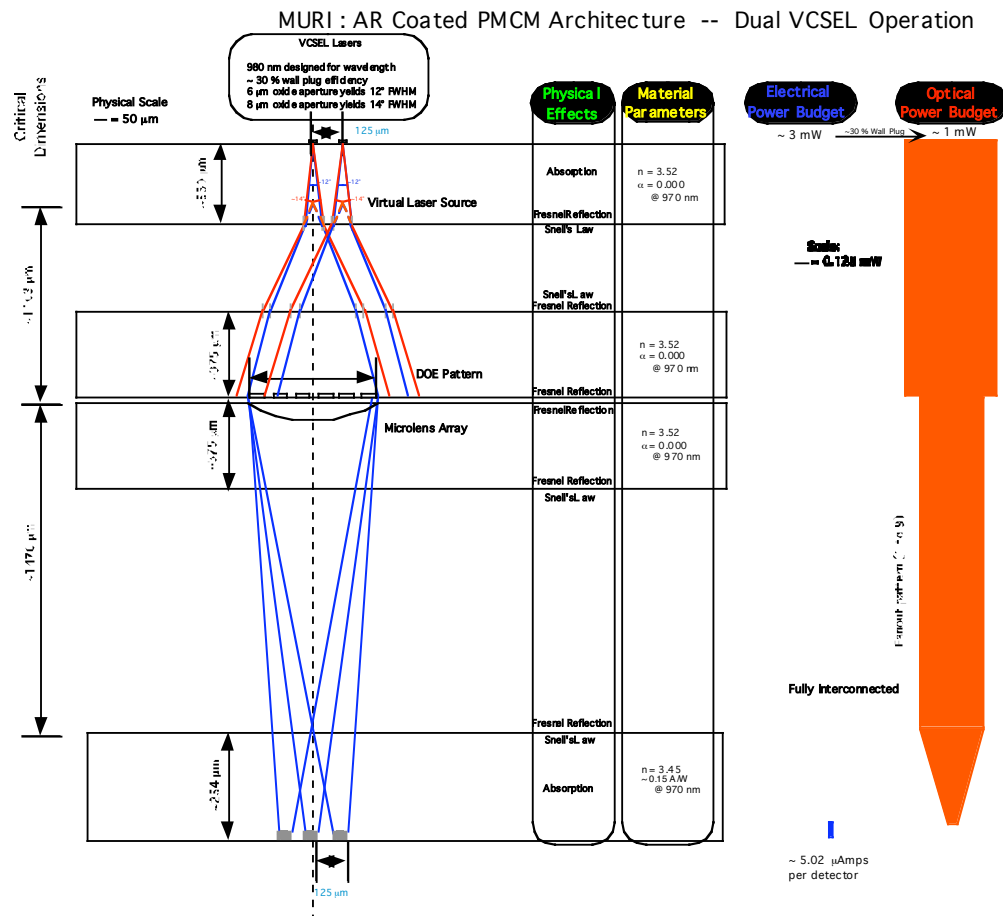
light spill into adjacent lenses, and any additional diffraction effects, are not considered here. The DOE and the lens array then cause light to be focused onto the back surface of the MOSIS-fabricated VLSI IC.

The relationship between the input optical intensity  $I_0$  and the optical intensity received on a single detector  $I_{det}$  is expressed by the following equation:

$$I_{det} = I_0 e^{-\alpha_{GaAs} d_{GaAs}} \eta_{GaAs/Air} \eta_{Air/GaAs} e^{-\alpha_{GaAs} d_{GaAs}} \chi_{DOE} \eta_{GaAs/Air} \eta_{Air/GaAs} e^{-\alpha_{GaAs} d_{Lens}} \eta_{GaAs/Air} \eta_{Air/Silicon} e^{-\alpha_{Si} d_{Si}} R_{Detector}$$

in which  $I_0$  is the input optical power,  $\alpha_{GaAs}$  is the absorption coefficient of GaAs,  $d_{GaAs}$  is the thickness of the VCSEL GaAs substrate,  $\eta_{GaAs/Air}$  is the Fresnel reflection coefficient of the GaAs/air interface,  $\eta_{Si/Air}$  is the Fresnel reflection coefficient of the silicon/air interface,  $d_{lens}$  is the thickness of the lens GaAs substrate,  $d_{si}$  is the thickness of the silicon substrate,  $\alpha_{si}$  is the absorption coefficient of silicon substrate,  $\chi_{DOE}$  is the percent of light diffracted into the DOE orders, and  $R_{detector}$  is the responsivity of the silicon detector. The surface normal transmittance through each interface is calculated by the expression  $T = 1 - R$ , and is equal to 68.6%, or a Fresnel reflection loss of 31.4% for a GaAs/Air interface. Due to the number of surfaces and their associated refractive indices, this number has a substantial effect on the final optical power reaching the silicon detectors. Assuming that 75% of the light is diffracted by the DOE into the necessary orders, indicated by the  $\chi_{DOE}$  term, and that a typical silicon detector's responsivity is 0.15 A/W, the final value of electrical current reaching the photodetector in our PMCM stack is only 0.22  $\mu$ A for a 1 mW input optical source, a value deemed too low to drive our VLSI electronics. A method to improve this situation is described in the next section.

When an appropriate AR coating is applied to layers in the PMCM, a significant increase in optical throughput is realized. Consider the diagram shown in Fig. 60 below. Assuming that 75% of the light is diffracted by the DOE into the necessary orders, indicated by the  $\chi_{DOE}$  term, and that the silicon detector's responsivity is 0.15 A/W, the final value of electrical current reaching the photodetector is 5.02  $\mu$ A for a 1 mW input optical source – more than enough to drive the VLSI computational electronics. This is in contrast to the previously calculated result of 0.22  $\mu$ A for a 1 mW input optical source. Essentially, we are removing the Fresnel reflections from six surfaces (five GaAs surfaces and the single back surface of the computational layer). Notice that the final power, *i.e.*, the power hitting the silicon photodetectors, is now only dependent on the input optical power, the efficiency of the DOE array, and the silicon VLSI substrate thickness. This makes any subsequent PMCM power optimization problems considerably easier.



**Fig. 60.** Schematic diagram of the PMCM containing AR coatings on all high refractive index surfaces, characterizing all unwanted optical losses. The width of the orange line represents the total optical power at that point in the PMCM stack.

## References

1. I. Biederman and G. Ju, "Surface *vs.* Edge-Based Determinants of Visual Recognition", *Cognitive Psychology*, **20**, 38-64, (1988).
2. I. Biederman, "Recognizing Depth-Rotated Objects: A Review of Recent Research and Theory", *Spatial Vision*, 241-253, (2000).
3. A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Boston, MA, (1987).
4. J. Fiser, S. Subramaniam, and I. Biederman, "Size Tuning in the Absence of Spatial Frequency Tuning in Object Recognition", *Vision Research*, **41**, 1931-1950, (2001).
5. K. Fukushima, S. Miyake, and T. Ito, "Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition", *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13**, 826-834, (1983).
6. S. Grossberg and E. Mignolla, "Neural Dynamics of Perceptual Grouping: Textures, Boundaries, and Emergent Segmentations", *Perception and Psychophysics*, **38**, 141-171, (1985).
7. M. H. Hayes, "The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of its Fourier Transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-30**, 140-154, (1982).
8. Y. Huang and B. K. Jenkins, "Non-Iterative Estimation with Perturbed Gaussian Markov Processes", In *Advances in Neural Information Processing Systems*, **Volume 18**, MIT Press, Boston, Massachusetts, (2006).
9. L. A. Iverson and S. W. Zucker, "Logical/Linear Operators for Image Curves", *IEEE PAMI*, **17**(10), 982-996, (1995).
10. B. K. Jenkins and A. R. Tanguay, Jr., "Photonic Implementations of Neural Networks", Ch. 9 in *Neural Networks for Signal Processing*, B. Kosko, (Ed.), Prentice-Hall, New Jersey, (1992), pp. 287-382.
11. M. Kapadia, M. Ito, C. D. Gilbert, and G. Westheimer, "Improvement in Visual Sensitivity by Changes in Local Context--Parallel Studies in Human Observers and in V1 of Alert Monkeys, *Neuron*, **15**, 843-856, (1995).
12. Z. Karim, "Thin Film Coatings for Optical Information Processing and Computing Applications", Ph.D. Thesis, University of Southern California, Los Angeles, CA, (December, 1993).
13. M. Lades, J. C. Vorbrueggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz, and W. Konen, "Distortion Invariant Object Recognition

in the Dynamic Link Architecture", IEEE Transactions on Computers, **42**, 300-311, (1993).

14. B. W. Mel, "SEEMORE: Combining Color, Shape, and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition", Neural Computation, **9**, 777-804, (1997).

15. M. Niederhouser, X. Yue, and I. Biederman, "Predicting Psychophysical Similarity of Complex Shapes from Measures of Physical Similarity", Journal of Vision, **6**(6), Abstract 320, <http://journalofvision.org/6/6/320/>, (2006).

16. K. Okada, S. Akamatsu and C. von der Malsburg, "Analysis and Synthesis of Pose Variations of Human Faces by a Linear PCMAP Model and its Application for Pose-Invariant Face Recognition System", in Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition, Grenoble, Switzerland, (March 26-30, 2000), IEEE Computer Society, pp. 142-149, (2000).

17. E. Peterhans and R. von der Heydt, "Mechanisms of Contour Perception in Monkey Visual Cortex. II. Contours Bridging Gaps", Journal of Neuroscience, **9**, 1749-1763, (1989).

18. G. Peters, B. Zitova, and C. von der Malsburg, "A Comparative Evaluation of Matching and Tracking Object Features for the Purpose of Estimating Similar-View-Areas of 3-Dimensional Objects, Technical Report IRINI 99-06, Institut für Neuroinformatik, Ruhr-Universität Bochum, (1999a).

19. G. Peters, B. Zitova, and C. von der Malsburg, "Two Methods for Comparing Different Views of the Same Object", Proceedings of the 10th British Machine Vision Conference (BMVC'99), Vol. 2, T. Pridmore and D. Elliman, (Eds.), University of Nottingham, pp. 493-502, (1999b).

20. G. Peters and C. von der Malsburg, "Interpolation of Novel Object Views from Two or Three Sample Views", accepted for presentation at ENAIS 2000 (Engineering Natural and Artificial Information Systems), Dubai, U. A. E., March 17-21, 2001, (2001a).

21. G. Peters and C. von der Malsburg, "View Reconstruction by Linear Combination of Sample Views", submitted to the Third International Conference on 3D Digital Imaging and Modeling (3DIM2001), Quebec, Canada, May 28-June 1, 2001, (2001b).

22. S. Subramaniam and I. Biederman, "Does Contrast Reversal Affect Object Identification?", in Investigative Ophthalmology and Visual Science, **38**, 999ff., The Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, (1997).

23. A. R. Tanguay, Jr. and B. K. Jenkins, "Modulator-Based Photonic Chip-to-Chip Interconnections for Dense Three-Dimensional Multichip Module Integration", U. S. Patent No. 5,568,574, (22 October, 1996).
24. J. Triesch and C. von der Malsburg, "Self-Organized Integration of Visual Cues for Face Tracking", Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition, Grenoble, France, March 28-30, 2000, IEEE Computer Society, pp. 102-107, (2000).
25. W. B. Veldkamp, "Wireless Focal Planes 'On the Road to Amacronic Sensors'", IEEE Jour. Quant. Electronics, **29**(2), 801-813, (1993).
26. C. von der Malsburg, "The Correlation Theory of Brain Function", Internal Report No. 81-2, Max Planck Institute for Biophysical Chemistry, Department of Neurobiology, Gottingen, Germany, (1981).
27. L. Shams and C. von der Malsburg, "The Role of Complex Cells in Object Recognition", Vision Research, **42**(22), 2547-2554, (2002).
28. J. Wieghardt and C. von der Malsburg, "Pose-Independent Object Representation by 2-D Views", Proceedings of the Conference on Biologically Motivated Computer Vision 2000, Seoul, Korea, (2000).
29. L. Wiskott, J.-M. Fellous, N. Krueger and C. von der Malsburg, "Face Recognition by Elastic Graph Matching", IEEE Trans. on Pattern Analysis and Machine Intelligence, **19**(7), 775-779, (1998).

## MURI Faculty Honors and Awards

These honors and awards were all received during the MURI research program, and recognize both the promise and accomplishments of key MURI faculty investigators.

### **Irving Biederman**

W. M. Keck Professor of Cognitive Neuroscience; Neuroscience Graduate Program, University of Southern California

- USC Associates Award for Creativity in Research and Scholarship (for the development of the Geon Theory of shape perception); March 9, 1999, Los Angeles, California; awarded by the University of Southern California.
- Mentor for Elizabeth Keiko Williams, Palos Verdes High School, Top 10 Finalist in the Intel Science Competition.

### **B. Keith Jenkins**

Professor of Electrical Engineering, University of Southern California

- Promoted to Professor of Electrical Engineering, September, 2004.

### **Bartlett W. Mel**

Associate Professor of Biomedical Engineering; Neuroscience Graduate Program, University of Southern California

- National Science Foundation Career Award, "Dendritic Subunits in Cortical Visual Processing and Development", 1998-2003.
- Promoted to Associate Professor of Biomedical Engineering, with tenure, May, 2000.

### **Christoph von der Malsburg**

Professor of Computer Science and Neurobiology; Neuroscience Graduate Program, University of Southern California, and Ruhr-Universität Bochum

- Karl-Heinz Beckurts Foundation Award for Outstanding Achievements in the Field of Neural Computing; December 4, 1998, Stuttgart, Germany; awarded by the Karl-Heinz Beckurts Foundation.
- The Koerber-Prize for Science in Europe 2000; September 7, 2000. The award also supports research over a three-year period to investigate elementary functions of the brain, specifically in the visual and olfactory systems, and to elucidate how the nervous systems of animals processes visual stimuli during recognition of an object, with the goal of applying such Gestalt perception to technology for the recognition of patterns and speech.



- Hebb Award, International Neural Network Society; 2003.

### **John O'Brien**

Professor of Electrical-Engineering-Electrophysics, University of Southern California

- 1998 Presidential Early Career Award for Scientists and Engineers; December, 1998, Washington, D.C. (awarded at a White House ceremony, February 10, 1999); awarded by the Office of Science and Technology Policy (OSTP) through the Department of Defense, Army Research Office (Army Research Office Young Investigator Program).
- 1998-99 Zumberge Fellow, July, 1998, Los Angeles, California; awarded by the University of Southern California.
- Promoted to Associate Professor of Electrical Engineering-Electrophysics, September, 2003.
- Promoted to Professor of Electrical Engineering-Electrophysics, January, 2006.

### **Bing Sheu**

Professor of Electrical Engineering-Electrophysics, Electrical Engineering Systems, and Biomedical Engineering, University of Southern California

- Distinguished Lecturer, IEEE Circuits and Systems Society, 1998-99
- Vice-President of Conferences, IEEE Circuits and Systems Society, 1998
- Editor-in-Chief, IEEE Transactions on VLSI Systems, 1998
- President-Elect, IEEE Circuits and Systems Society, 1999
- Editor-in-Chief, IEEE Transactions on Multimedia, 1999

### **Armand R. Tanguay, Jr.**

Professor of Electrical Engineering-Electrophysics, Chemical Engineering and Materials Science, and Biomedical Engineering; Neuroscience Graduate Program, University of Southern California

- Fellow, American Association for the Advancement of Science (AAAS), November, 1999; awarded in Washington, D.C. at the AAAS Annual Meeting 2000, February 19, 2000. Citation: "for distinguished contributions to physical optics, optical materials and devices, and optical information processing and computing, including the invention of stratified volume holographic optical elements".

- Faculty Fellow, Center for Excellence in Teaching, University of Southern California, 2001-2005.
- Promoted to Professor of Electrical Engineering-Electrophysics, Materials Science, and Biomedical Engineering, January, 2001.
- Distinguished Faculty Fellow, Center for Excellence in Teaching, University of Southern California, 2005-present.
- Teacher of the Year, 2002, Latter Day Saints Student Association, University of Southern California.

## **Scientific Personnel**

### **Key Faculty Investigators**

The eight key faculty members involved in the MURI research program on Adaptive Optoelectronic Eyes at USC are listed below, along with areas of research expertise and interest. This list illustrates the interdisciplinary contributions of each faculty member to the integrated effort.

Professor Madhukar was not funded by the MURI effort, but contributed nonetheless in the areas indicated, as well as through the AFOSR MURI effort on multiple quantum well and quantum box infrared (IR) sensors.

**Prof. Irving Biederman**

**William T. Keck Professor**

**Psychology**

**Member, Neuroscience Graduate Program**

Psychology of Vision; Experimental Tests of Human Visual Capabilities; Development of Higher-Level Models and Vision Algorithms; Development of Geon Theory of Vision

**Prof. Christoph von der Malsburg**

**Computer Science and Neurobiology**

**Member, Neuroscience Graduate Program**

Physiology and Psychology of Vision; Development of Low-Level and Mid-Level Vision Algorithms Based on Spatial Relationships and Feature Similarities; Mapping of Low-Level and Mid-Level Vision Algorithms; Face Recognition; Image Reconstruction

**Prof. Bartlett Mel**

**Biomedical Engineering**

**Member, Neuroscience Graduate Program**

Development of Low-Level and Mid-Level Vision Algorithms Based on Co-Occurrences of Extracted Features; Mapping of Low-Level and Mid-Level Vision Algorithms; Testing of Vision Algorithms in Realistic Environments

**Prof. B. Keith Jenkins**

**Electrical Engineering-Systems**

Mapping of Vision Algorithms into Electronic/Photonic Hardware Implementations; DOE and Optical Systems Design

**Prof. Armand R. Tanguay, Jr.**  
**Electrical Engineering-Electrophysics, Materials Science, and Biomedical Engineering**

**Member, Neuroscience Graduate Program**

Hybrid Analog/Digital VLSI Design; Diffractive Optical Element Fabrication and Testing; Stratified Volume Holographic Optical Elements; Stratified Volume Diffractive Optical Elements; Integrated Optical Devices (Optical Power Bus); Flip-Chip Bonding and Device Packaging

**Prof. Bing Sheu**

**Electrical Engineering-Electrophysics, Electrical Engineering-Systems, and Biomedical Engineering**

Hybrid Analog/Digital VLSI Design; Cellular Neural Network Designs; VLSI Chip Testing and Analysis; Active Pixel CMOS Sensor Arrays

**Prof. John O'Brien**

**Electrical Engineering-Electrophysics**

E-Beam Lithography for Diffractive Optical Element Fabrication; Nanofabrication Technology; Vertical Cavity Surface Emitting Laser Arrays

**Prof. Anupam Madhukar**

**Kenneth T. Norris Professor**

**Materials Science and Physics**

Molecular Beam Epitaxy (MBE) Growth of Multiple Quantum Well (MQW) Modulator, Detector, and Vertical Cavity Surface Emitting Laser Arrays; Nanofabrication Technology; Focused Ion Beam Fabrication of Diffractive Optical Elements; Quantum Dot IR Photodetector Arrays

### **Affiliated Faculty Investigators**

The four faculty members at both USC and other universities that became involved in the MURI research program on Adaptive Optoelectronic Eyes at USC since its inception are listed below, along with their areas of research expertise and interest. This list illustrates the complementary contributions of these faculty members to the overall research program.

**Prof. P. Daniel Dapkus**

**University of Southern California**

Metal-Organic Chemical Vapor Deposition (MOCVD) Growth of Multiple Quantum Well (MQW) Detector and Vertical Cavity Surface Emitting Laser Arrays; Design and Fabrication of Ultra-Low Threshold Vertical Cavity Surface Emitting Laser Arrays

**Prof. Nicholas George**

**University of Rochester, Institute of Optics**

Smart Cameras; Diffractive Optical Elements; Stratified Volume Holographic Optical Elements; Metrics for Automatic Evaluation of Image Quality

**Prof. Gregory P. Nordin**

**University of Alabama, Huntsville**

Diffractive Optical Element Design, Fabrication, and Testing; Rigorous Coupled Wave Analysis; Stratified Volume Holographic Optical Elements; Stratified Volume Diffractive Optical Elements

**Prof. Mandyam Srinivasan**

**Research School of Biological Sciences**

**The Institute of Advanced Studies**

**Australian National University**

**Canberra, Australia**

Insect Visual Systems, Visual Pursuit and Navigation Algorithms, Biologically-Inspired Artificial Vision Systems

**MURI Postdoctoral Fellows, Graduate Research Assistants,  
Undergraduate Research Assistants, and Administrative Staff**

**Postdoctoral Fellows**

Dr. Gary Holt (Related research)

Dr. Patrick Nasiatka (2003-2004)

**Graduate Research Assistants**

Kevin Archie

Neil Abbasi (Related research)

Moshe Bar (Related research)

Ran Carmi

Jaeyoun Cho

C. Eckes (Related research)

E. Elagin (Related research)

Hsing-Hua Fan

Hung-Min Jen

Hai Hong  
Yunsong Huang  
Andrea Kosta  
Po-Tsung Lee  
H. Loos (Related research)  
Jaw-Chyng (Lormen) Lue  
J. Luecke (Related research)  
Michael Mangini  
T. Maurer (Related research)  
Patrick Nasiatka (1998-2003)  
H. Neven (Related research)  
Kazunori Okada  
Panayiota Poirazi  
C. Prodoehl (Related research)  
Roshanak Shafiiha  
Ladan Shams (Related research)  
J. Steffens (Related research)  
Suresh Subramaniam (Related research)  
Nan-Kyung Suh (Related research)  
J. Triesch (Related research)  
Edward Vessel  
L. Wiskott (Related research)  
I. Wundrich (Related research)  
Joshua Wyner (1999-2000)

Junmei Zhu

**Undergraduate Research Assistants**

Joshua Wyner (1998-99)

**Administrative Staff**

Gloria Halfacre

Karen Johnson

**Degrees Conferred**

JoszeF Fiser, Ph.D., Neuroscience, University of Southern California, (August, 1998).

Moshe Bar, Ph.D., Psychology, University of Southern California, (December, 1998).

Suresh Subramaniam, Ph.D., Psychology, University of Southern California, (December, 1998).

Ingo Wundrich, Diploma, Information Technology, Ruhr-Universität Bochum, (December, 1998).

Triesch, Jochen, Ph.D., Physics and Astronomy, Ruhr-Universität Bochum, (July, 1999).

Ladan Shams, Ph.D., Computer Science, University of Southern California, (August, 1999).

Kai Bruenenberg, Diploma, Physics and Astronomy, Ruhr-Universität Bochum, (August, 1999).

Thomas Maurer, Ph.D., Physics and Astronomy, Ruhr-Universität Bochum, (November, 1999).

Michael Poetzsch, Ph.D., Physics and Astronomy, Ruhr-Universität Bochum, (December, 1999).

Achim Schaefer, Diploma, Physics and Astronomy, Ruhr-Universität Bochum, (February, 2000).

Peter Kalocsai, Ph.D., Psychology, University of Southern California, (August, 2000).

Yiota Poirazi, Ph.D., Biomedical Engineering, University of Southern California, (August, 2000).

Hai Hong, Ph.D., Computer Science, University of Southern California, (December, 2000).

Gabriele Peters, Ph.D., Computer Science, Universität Bielefeld, (April, 2001).

Jan Wieghardt, Ph.D., Physics, Ruhr-Universität Bochum, (July, 2001).

Marissa Nederhouser, Ph.D., Psychology, University of Southern California, (June, 2002).

Hartmut S. Loos, Ph.D., Computer Science, Universität Bielefeld, (November, 2002).

Patrick Nasiatka, Ph.D., Electrical Engineering-Electrophysics, University of Southern California, (June, 2003).

Junmei Zhu, Ph.D., Computer Science, University of Southern California, (July, 2003).

Po-Tsung Lee, Ph.D., Electrical Engineering, University of Southern California, (2003).

Michael Mangini, Ph.D., Psychology, University of Southern California, (December, 2003).

Edward Vessel, Ph.D., Neuroscience, University of Southern California, (August, 2004).

Xiangyu Tang, Ph.D., Computer Science, University of Southern California, (September, 2004).

Roshanak Shafiiha, M.S., Ph.D., Electrical Engineering-Electrophysics, University of Southern California, (2000; 2004).

Nan-Kyung Suh, Ph.D., Electrical Engineering, University of Southern California, (January, 2005).

Carsten Prodoehl, Ph.D., Biology, Universität Bielefeld, (January, 2005).

Chunhong Zhou, Ph.D., Biomedical Engineering, University of Southern California, (January, 2005).

Andreas Tewes, Ph.D., Physics, Ruhr-Universität Bochum, (February, 2006).

Neil Abbasi, M.S., Engr., Electrical Engineering, University of Southern California, (May, 2006).



## List of Publications

### Book Chapters

1. I. Biederman and P. Kalocsai, "Neural and Psychophysical Analysis of Object and Face Recognition", in *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. Huang (Eds.), NATO ASI Series F, Springer-Verlag, (1998), pp. 3-25.
2. P. Kalocsai and I. Biederman, "Differences of Face and Object Recognition in Utilizing Early Visual Information", in *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. Huang (Eds.), NATO ASI Series F, Springer-Verlag, (1998), pp. 492-502.
3. K. Okada, J. Steffens, T. Maurer, Hai Hong, E Elagin, H. Neven and C. von der Malsburg, "The Bochum/USC Face Recognition System and How it Fared in the FERET Phase III Test", in *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. S. Huang (Eds.), NATO ASI Series F, Springer-Verlag, (1998), pp. 186-205.
4. B. W. Mel, "Why Have Dendrites? A Computational Perspective", in *Dendrites*, G. Stuart, N. Spruston, and M. Hausser, (Eds.), Oxford University Press, (1999), pp. 271-289.
5. B. K. Jenkins and A. R. Tanguay, Jr., "Photonic Implementations of Neurobiologically Inspired Networks," in M. Arbib, Ed., *Handbook of Brain Theory and Neural Networks, Second Edition*, MIT Press, (2002), pp. 889-893.
6. A. R. Tanguay, Jr. and B. K. Jenkins, "Hybrid Electronic/Photonic Multichip Modules for Vision and Neural Prosthetic Applications," Ch. 14 in T. W. Berger and D. L. Glanzman, Eds, *Toward Replacement Parts for the Brain: Implantable Biomimetic Electronics as the Next Era in Neural Prosthetics*, MIT Press, Cambridge, (2005), pp. 295-334.
7. L. Wiskott, C. von der Malsburg, and A. Weitzenfeld, "Face Recognition by Dynamic Link Matching", in *The Neural Simulation Language: A System for Brain Modeling*, A. Weitzenfeld, M. A. Arbib, and A. Alexander, Eds., MIT Press, Cambridge, (2002), ISBN 0-262-73149-5, Chapter 18, pp. 343-372.
8. H. Loos and C. von der Malsburg, "1-Click Learning of Object Models for Recognition", in *Biologically Motivated Computer Vision 2002 (BMCV 2002, Tuebingen, Germany)*, H. H. Buelthoff, S.-W. Lee, T. A. Poggio, and C. Wallraven, Eds., Lecture Notes in Computer Science 2525, Springer Verlag, Heidelberg, (2002), pp. 377-386.
9. C. von der Malsburg, "Dynamic Link Architecture" in *The Handbook of*

*Brain Theory and Neural Networks*, Second Edition, M. Arbib, Ed., MIT Press, (2002), pp. 365-368.

10. C. von der Malsburg, "Self-Organization in the Brain", in *The Handbook of Brain Theory and Neural Networks*, Second Edition, M. Arbib, Ed., MIT Press, (2002), pp. 1002-1005.

11. J. O'Brien and W. Kuang, "Diffractive Systems: Photonic Crystal Lasers, Cavities, and Waveguides", Invited Contribution to the *Encyclopedia of Modern Optics*, Academic Press, (2004).

12. J. O'Brien, W. Kuang, P.-T. Lee, J. R. Cao, and C. Kim, "Photonic Crystal Lasers", Invited Contribution to the *Encyclopedia of Nanoscience and Nanotechnology*, American Scientific Publishers, (2004).

### **Journal Publications**

1. A. E. Bond, P. D. Dapkus, and J. D. O'Brien, "Aperture Placement Effects in Oxide-Defined Vertical-Cavity Surface-Emitting Lasers", *IEEE Photonics Technology Letters*, **10**, 1362-1364, (1998).

2. M. Bar and I. Biederman, "Subliminal Visual Priming", *Psychological Science*, **9**, 464-469, (1998).

3. A. E. Bond, P. D. Dapkus, and J. D. O'Brien, "Aperture Dependent Loss Analysis in Vertical-Cavity Surface-Emitting Lasers", *IEEE Photonics Technology Letters*, **11**, 397-399, (1999).

4. D. V. Tishinin, P. D. Dapkus, A. E. Bond, I. Kim, C. K. Lin, and J. D. O'Brien, "Vertical Resonant Couplers with Precise Coupling Efficiency Control Fabricated by Wafer Bonding", *IEEE Photonics Technology Letters*, **11**, 1003-1005, (1999).

5. A. E. Bond, P. D. Dapkus, and J. D. O'Brien, "Design of Low-Loss Single-Mode Vertical-Cavity Surface-Emitting Lasers", *IEEE Journal Selected Topics in Quantum Electronics*, **5**, 574-581, (1999).

6. O. J. Painter, R. K. Lee, A. Scherer, A. Yariv, J. D. O'Brien, P. D. Dapkus, and I. Kim, "Two-Dimensional Photonic Band-Gap Defect Mode Laser", *Science*, 1819-1821, (1999).

7. L. Shams and C. von der Malsburg, "Are Object Shape Primitives Learnable?", *Journal of NeuroComputing*, **26-27**, 855-863, (1999).

8. O. J. Painter, A. Husain, A. Scherer, J. D. O'Brien, I. Kim, and P. D. Dapkus, "Room Temperature Photonic Crystal Defect Lasers at Near-Infrared Wavelengths in InGaAsP", *IEEE Journal Lightwave Technology*, 2082-2087, (1999).

9. "Two-Dimensional Photonic Bandgap Defect Laser", *Optics and Photonics News*, September, 1999, p. 42, (1999).
10. I. Biederman and M. Bar, "One-Shot Viewpoint Invariance in Matching Novel Objects, *Vision Research*, **39**, 2885-2899, (1999).
11. "Photonic Crystal Nanocavity Lasers", *Optics and Photonics News*, December, 1999, 21-22, (1999).
12. "Lasing Demonstrated in Tiny Cavities Made with Photonic Crystals", *Physics Today*, 20-22, September, (1999).
13. S. Piazzolla and B. K. Jenkins, "Dynamics During Holographic Exposure in Photopolymers for Single and Multiplexed Gratings," *Journal of Modern Optics*, **46**(14), 2079-2110, (1999).
14. B. W. Mel, "Think Positive to Find Parts", *Nature*, **401**, 759-760, (1999).
15. B. W. Mel and J. Fiser, "Minimizing Binding Errors Using Learned Conjunctive Features", *Neural Computation*, **12**, 247-278, (2000).
16. P. Poirazi and B. W. Mel, "Choice and Value Flexibility Jointly Contribute to the Capacity of a Subsampled Quadratic Classifier", *Neural Computation*, **12**, 1189-1205, (2000).
17. K. A. Archie and B. W. Mel, "A Model for Intradendritic Computation of Binocular Disparity, *Nature Neuroscience*, **3**, 54-63, (2000).
18. S. Piazzolla and B. K. Jenkins, "First Harmonic Diffusion Model for Holographic Grating Formation in Photopolymers," *Journal of the Optical Society of America B*, **17**(7), 1147-1157, (2000).
19. I. Biederman, "Recognizing Depth-Rotated Objects: A Review of Recent Research and Theory", *Spatial Vision*, 241-253, (2000).
20. O. Painter, A. Husain, A. Scherer, P. Lee, I. Kim, J. D. O'Brien, and P. D. Dapkus, "Lithographic Tuning of a Two-Dimensional Photonic Crystal Laser Array", *IEEE Photonics Technology Letters*, **12**, 1126-1128, (2000).
21. P. Kalocsai, C. von der Malsburg, and J. Horn, "Face recognition by statistical analysis of feature detectors, *Image And Vision Computing*, **18**(4), 273-278, (2000).
22. A. Scherer, O. Painter, A. Husain, J. Vuckovic, P. D. Dapkus, and J. O'Brien, "Photonic Crystal Nanocavity Lasers", *International Journal of High Speed Electronics and Systems (Singapore)*, **10**(1), 387-391, (2000).
23. I. Biederman and M. Bar, "Views on Views: Response to Hayward and Tar (2000)", *Vision Research*, **40**, 3901-3905, (2000).

24. J. Fiser and I. Biederman, "Invariance of Long-Term Visual Priming to Scale, Reflection, Translation, and Hemisphere", *Vision Research*, **41**, 221-234, (2001).
25. J. Fiser, S. Subramaniam, and I. Biederman, "Size Tuning in the Absence of Spatial Frequency Tuning in Object Recognition", *Vision Research*, **41**, 1931-1950, (2001).
26. R. Vogels, I. Biederman, M. Bar, and A. Lorincz, "Inferior Temporal Neurons Show Greater Sensitivity to Nonaccidental than Metric Differences", *Journal of Cognitive Neuroscience*, **13**, 444-453, (2001).
27. O. J. Painter, K. Srinivasan, J. D. O'Brien, A. Scherer, and P. D. Dapkus, "Tailoring of the Resonant Mode Properties in Two-Dimensional Photonic Crystal Slab Waveguides," *Journal of Optics A: Pure and Applied Optics*, **3**, S161-S170, (2001).
28. J. Triesch and C. von der Malsburg: "A System for Person-Independent Hand Posture Recognition Against Complex Backgrounds", *PAMI*, **23**(12), 1449-1453, (2001).
29. A. Lőrincz, G. Szirtes, B. Takács, I. Biederman, and R. Vogels, "Relating Priming and Repetition Suppression", *International Journal of Neural Systems*, **12**, 187-202, (2002).
30. J. R. Cao, P.-T. Lee, S.-J. Choi, R. Shafiiha, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Nanofabrication of Photonic Crystal Membrane Lasers," *Journal of Vacuum Science and Technology B*, **20**, 618-621, (2002).
31. L. Shams and C. von der Malsburg, "Acquisition of Visual Shape Primitives", *Vision Research*, **42**(17), 2105-2122, (2002).
32. P.-T. Lee, J.-R. Cao, S.-J. Choi, Z.-J. Wei, J. D. O'Brien, and P. D. Dapkus, "Room Temperature Operation of VCSEL-Pumped Photonic Crystal Lasers," *IEEE Photonics Technology Letters*, **14**, 435-437, (2002).
33. C. Kim, W. J. Kim, A. Stapleton, J. R. Cao, J. D. O'Brien, and P. D. Dapkus, "Quality Factors in Single Defect Photonic Crystal Lasers with Asymmetric Cladding Layers," *Journal of the Optical Society of America B*, **19**, 1777-1781, (2002).
34. L. Shams and C. von der Malsburg, "The Role of Complex Cells in Object Recognition", *Vision Research*, **42**(22), 2547-2554, (2002).
35. P.-T. Lee, J.-R. Cao, S.-J. Choi, Z.-J. Wei, J. D. O'Brien, and P. D. Dapkus, "Operation of Photonic Crystal Membrane Lasers Above Room Temperature," *Applied Physics Letters*, **81**, 3311-3313, (2002).

36. J. Triesch and C. von der Malsburg, "Classification of Hand Postures Against Complex Backgrounds Using Elastic Graph Matching", *Image and Vision Computing*, **20**(13-14), 937-943, (2002).
37. J.-R. Cao, P.-T. Lee, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Lithographic Fine Tuning of VCSEL Pumped 2-D Photonic Crystal Lasers," *Journal of Nanoscience and Nanotechnology*, **2**, 313-315, (2002).
38. J. Zhu and C. von der Malsburg, "Synapto-Synaptic Interactions Speed up Dynamic Link Matching", *Neurocomputing*, **44**, 721-728, (2002).
39. C. Zhou, G. Holt, and B. Mel, "Automatic Line-Drawings Extraction from Complex Scenes", [Abstract], *Journal of Vision*, **2**(7), 501a, <http://journalofvision.org/2/7/501/>, doi:10.1167/2.7.501 (2002).
40. C. Prodoehl, R. P. Würtz, and C. von der Malsburg, "Learning the Gestalt Rule of Collinearity from Object Motion", *Neural Computation*, **15**, 1865-1896, (2003).
41. J.-R. Cao, W. Kuang, S.-J. Choi, P.-T. Lee, J. D. O'Brien, and P. D. Dapkus, "Threshold Dependence on the Spectral Alignment Between the Quantum-Well Gain Peak and the Cavity Resonance in InGaAsP Photonic Crystal Lasers," *Applied Physics Letters*, **83**, 4107-4109, (2003).
42. J. Zhu and C. von der Malsburg, "Object Recognition by Dynamic Link Matching in Biologically Realistic Time", [Abstract], *Journal of Vision*, **3**(9), 195a, (2003).
43. J. Zhu and C. von der Malsburg, "Learning Control Units for Invariant Recognition", *Neurocomputing*, **52**, 447-453, (2003).
44. C. Zhou and B. W. Mel, "Combining Multiple Cues for Contour Detection: Lessons from (and to) the Visual Cortex", [Abstract], *Journal of Vision*, **3**(9), 252a, <http://journalofvision.org/3/9/252/>, doi:10.1167/3.9.252, (2003).
45. J. Luecke, and C. von der Malsburg, "Rapid Processing and Unsupervised Learning in a Model of the Cortical Macrocolum", *Neural Computation*, **16**, 501-533, (2004).
46. J. Zhu and C. von der Malsburg, "Maplets for Correspondence-Based Object Recognition", *Neural Networks*, **17**, 1311-1326, (2004).
47. I. J. Wundrich, C. von der Malsburg and R. P. Würtz, "Image Representation by Complex Cell Responses", *Neural Computation*, **16**, 2563-2575, (2004).
48. C. Zhou and B. W. Mel, "Combining Cues for Boundary Detection Using the "Mixture of Specialists" Model", [Abstract], *Journal of Vision*, **5**(8), 501a, <http://journalofvision.org/5/8/501/>, doi:10.1167/5.8.501, (2005).

49. J.-R. Cao, W. Kuang, Z.-J. Wei, S.-J. Choi, H. Yu, M. Bagheri, J. D. O'Brien, and P. D. Dapkus, "Sapphire-Bonded Photonic Crystal Microcavity Lasers and Their Far-Field Radiation Patterns," *IEEE Photonics Technology Letters*, **17**(1), 4-6, (2005).

50. W. Kuang, J. R. Cao, T. Yang, S.-J. Choi, P.-T. Lee, J. D. O'Brien, and P. D. Dapkus, "Classification of Modes in Suspended Membrane 19-Missing Hole Photonic Crystal Microcavities," *Journal of the Optical Society of America B*, **22**, 1092-1099, (2005).

51. W. Kuang, J. R. Cao, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Modified Suspended Membrane Photonic Crystal  $D_3$  Laser Cavity with Improved Side Mode Suppression Ratio," *Photonics Technology Letters*, **17**(5), 941-942, (2005).

52. C. Eckes, J. Triesch and C. von der Malsburg, "Analysis of Cluttered Scenes Using an Elastic Matching Approach for Stereo Images", *Neural Computation*, **18**, 1441-1471, (2006).

53. J. Zhu and C. von der Malsburg, "Associative Memory of Connectivity Patterns", *Neurocomputing*, (in print, 2006).

#### **Conference Presentations (Published Proceedings)**

1. C. von der Malsburg, L. Shams, and U. Eysel, "Recognition of Images From Complex Cell Responses", in *Proceedings of the Society for Neuroscience Meeting*, Los Angeles, California, (November 7-12, 1998), pg. 261, (1998).

2. G. Peters, B. Zitova, and C. von der Malsburg, "A Comparative Evaluation of Matching and Tracking Object Features for the Purpose of Estimating Similar-View-Areas of 3-Dimensional Objects", *Technical Report IRINI 99-06*, Institut für Neuroinformatik, Ruhr-Universität Bochum, (1999a).

3. G. Peters, B. Zitova, and C. von der Malsburg, "Two Methods for Comparing Different Views of the Same Object", in *Proceedings of the 10th British Machine Vision Conference (BMVC'99)*, Vol. 2, T. Pridmore and D. Elliman, Eds., University of Nottingham, pp. 493-502, (1999b).

4. J. Triesch and C. von der Malsburg, "Self-Organized Integration of Visual Cues for Face Tracking", in *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, (March 26-30, 2000), *IEEE Computer Society*, pp. 102-107, (2000).

5. K. Okada, S. Akamatsu and C. von der Malsburg, "Analysis and Synthesis of Pose Variations of Human Faces by a Linear PCMAP Model and its Application for Pose-Invariant Face Recognition System", in *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition*,

Grenoble, Switzerland, (March 26-30, 2000), IEEE Computer Society, pp. 142-149, (2000).

6. C. von der Malsburg, K. Reiser, G. Peters, J. Wiegardt, and K. Okada, "3-D Object Representation by 2-D Views", in Proceedings of the 6th ATR Symposium on Face and Object Recognition, Kyoto, Japan, pp. 11-12, (1999).

7. R. Vogels, I. Biederman, and M. Bar, "Sensitivity of Macaque Temporal Neurons to Differences in View-Invariant *vs.* Metric Properties of Depth-Rotated Objects" Meetings of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May, 1999, in Investigative Ophthalmology & Visual Science, **40**, S776, (1999).

8. E. A. Vessel, S. Subramaniam, and I. Biederman, "A Change in Contrast Polarity at an L-Junction Unbinds its Segments", Meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May, 1999, in Investigative Ophthalmology & Visual Science, **40**, 810, (1999).

9. M. C. Mangini, I. Biederman, and E. Williams, "The Effect of Test-Context Junction Discontinuities in Perceived Lightness", Meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May, 1999, in Investigative Ophthalmology & Visual Science, **40**, 747, (1999).

10. J. Wiegardt and C. von der Malsburg, "Pose-Independent Object Representation by 2-D Views", Proceedings of the Conference on Biologically Motivated Computer Vision 2000, Seoul, Korea, (2000).

11. M. C. Mangini, I. Biederman, and A. Kosta, "Is Greater Accuracy for Gender than Person Discrimination of Faces a Consequence of Class Uncertainty? Evidence from Normals and a Prosopagnosic", Meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May, 2000, in Investigative Ophthalmology & Visual Science, **41**, 225, (2000).

12. E. A. Vessel and I. Biederman, "Brightness Judgments Within Minimal Part Types are Easier than Between Part Types", Meeting of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, Florida, May, 2000, in Investigative Ophthalmology & Visual Science, **41**, 226, (2000).

13. C. von der Malsburg and J. Zhu, "Fast Dynamic Link Matching by Communicating Synapses", Proceedings of the 7th Joint Symposium on Neural Computation, University of Southern California, 122-129, (2000).

14. A. R. Tanguay, Jr., B. K. Jenkins, C. von der Malsburg, B. Mel, G. Holt, J. O'Brien, I. Biederman, A. Madhukar, P. Nasiatka, and Y. Huang, "Vertically Integrated Photonic Multichip Module Architecture for Vision Applications", Proceedings of the International Conference on Optics in Computing (OC 2000), Quebec City, Canada, (June 18-23, 2000), in Proceedings of the Society for Photo-Optical Instrumentation Engineers, **4089**, 584-600, (2000).

15. G. R. Holt and B. W. Mel, "Unsupervised Learning of the Nonclassical Surround", submitted to the Neural Information Processing Systems Conference, Denver, Colorado, (2000).
16. G. Peters and C. von der Malsburg, "Interpolation of Novel Object Views from Two or Three Sample Views", ENAIS 2000 (Engineering Natural and Artificial Information Systems), Dubai, U. A. E., March 17-21, 2001, (2001).
17. K. A. Archie and B. W. Mel, "Dendritic Compartmentalization Could Underlie Competition and Attentional Biasing of Simultaneous Visual Stimuli", in Advances in Neural Information Processing Systems (NIPS), Vol. 13, T. K. Leen, T. G. Dietterich, and V. Tresp, (Eds.), MIT Press, Cambridge, MA, 82-88, (2001).
18. G. Peters and C. von der Malsburg, "View Reconstruction by Linear Combination of Sample Views", Third International Conference on 3D Digital Imaging and Modeling (3DIM2001), Quebec, Canada, May 28-June 1, 2001, (2001).
19. J. Luecke, C. von der Malsburg, and R. P. Würtz, "Macrocolumns as Decision Units", Proceedings of ICANN 2002, Springer, pp. 57-62, (2002).
20. N. Serrano, A. Ortega, S. Wu, K. Okada, and C. von der Malsburg, "Compression for Distributed Face Recognition", in Proceedings of the Multimodal User Authentication Workshop, Santa Barbara, CA, (December 11-12, 2003), <http://mmua.cs.ucsb.edu>.
21. A. R. Tanguay, Jr., B. K. Jenkins, C. von der Malsburg, B. Mel, I. Biederman, J. O'Brien, and A. Madhukar, "Vertically-Integrated Photonic Multichip Module for Vision Applications", Proceedings of the 2004 IEEE International Joint Conference on Neural Networks, 25-29 July, 2004, 3, 1749, (2004).
22. Y. Huang and B. Keith Jenkins, "Non-Iterative Estimation with Perturbed Gaussian Markov Processes," Nineteenth Annual Conference on Neural Information Processing Systems, Vancouver, B.C., December, 2005, Paper 621; Advances in Neural Information Processing Systems (Proceedings), Vol. 18, MIT Press, (2006).

### **Conference and Technical Presentations**

1. A. R. Tanguay, Jr., "Photonic Implementations of Neural Networks", Symposium on Hardware Implementations of Neural Networks, Institut für Neuroinformatik Systembiophysik, Ruhr-Universität Bochum, Bochum, Germany, (June 30, 1998); (Invited Presentation).
2. A. R. Tanguay, Jr., "Parallel Neural Probe Arrays for Hippocampal Slice Culture and Dissociated Cell Growth Neuron-Silicon Interface Devices", Symposium on Hardware Implementations of Neural Networks, Institut für



Neuroinformatik Systembiophysik, Ruhr-Universität Bochum, Bochum, Germany, (July 2, 1998); (Invited Presentation).

3. A. R. Tanguay, Jr., "Technologies for the Implementation of Neural Networks", Neuroinformatik Seminar, Institut für Neuroinformatik Systembiophysik, Ruhr-Universität Bochum, Bochum, Germany, (July 2, 1998); (Invited Presentation).

4. B. K. Jenkins, "Information Display and Driver Interface in the Automotive Cockpit," TRW Technology/Requirements Interchange Symposium, Automotive Electronics/Space and Defense, Marriott Hotel, Redondo Beach, Calif., (July 8-9, 1998).

5. A. R. Tanguay, Jr., "Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures", Kick-Off Meeting for Multidisciplinary University Research Initiative (MURI), Program on Adaptive Optoelectronic Eye, Ann Arbor, Michigan, (July 11-14, 1998).

6. C. von der Malsburg, "Biologically-Inspired Algorithms for Adaptive Visual Object Recognition", Kick-Off Meeting for Multidisciplinary University Research Initiative (MURI), Program on Adaptive Optoelectronic Eye, Ann Arbor, Michigan, (July 11-14, 1998).

7. B. K. Jenkins, "From Vision Algorithms to Hybrid Electronic/Photonic Hardware Implementations", Kick-Off Meeting for Multidisciplinary University Research Initiative (MURI), Program on Adaptive Optoelectronic Eye, Ann Arbor, Michigan, (July 11-14, 1998).

8. A. R. Tanguay, Jr., "Emerging Smart Camera Technologies: Toward an Adaptive Optoelectronic Eye", Eastman Kodak Company, Kodak Research Laboratories, Rochester, New York, (July 22, 1998).

9. I. Biederman, "The Neurocomputational Basis of Face and Object Recognition" Stockholm Workshop on Computational Vision, Rosenön, Sweden, (August 4-7, 1998).

10. A. R. Tanguay, Jr., "Emerging Smart Camera Technologies: Toward an Adaptive Optoelectronic Eye", Teledyne Electronic Technology R&D Symposium, Los Angeles, California, (September 1-2, 1998).

11. M. Baudry, T. W. Berger, R. D. Brinton, B. Sheu, and A. R. Tanguay, Jr., "Hybrid Biological-Electronic Biosensor for Detection of Chemically- or Biologically-Induced Cognitive Impairment", DARPA Tissue Based Biosensors Program Annual Program Review, Washington, D.C., (October 12-15, 1998).

12. A. R. Tanguay, Jr., "Photonic Implementations of Neural Networks", Biomedical Simulations Resource (BMSR) Workshop, (held in Marina del Rey,

CA), University of Southern California, Los Angeles, California, (November 14, 1998).

13. I. Biederman and M. Bar, "Cortical Localization of Subliminal Visual Priming, Annual Meeting of the Psychonomic Society, Dallas, Texas, (November, 1998).

14. J. J. Peissig, M. E. Young, Jr., E. A. Wasserman, and I. Biederman, "The Pigeon's Discrimination of Single Geons", Annual Meeting of the Psychonomic Society, Dallas, Texas, (November, 1998).

15. M. C. Mangini, I. Biederman, and E. K. Williams, "Perceived Lightness as a Measure of Perceptual Grouping", Annual Object Perception and Memory Meeting, Dallas, Texas, (November, 1998).

16. E. Vessel, S. Subramaniam, and I. Biederman, "When Does Variation in Contrast Polarity Affect Contour Grouping in Object Recognition?", Annual Object Perception and Memory Meeting, Dallas, Texas, (November, 1998).

17. G. Sáry, G. Kovács, K. Köteles, G. Benedek, J. Fiser, and I. Biederman, "Selectivity Variations in Monkey Inferior Temporal Neurons for Intact and Contour-Deleted Line Drawings", Annual Meeting of the Society for Neuroscience, Los Angeles, California, (November, 1998).

18. M. Bar and I. Biederman, "Subliminal Visual Priming Transfers Within but not Between Visual Quadrants", Meetings of the Society for Neuroscience, Los Angeles, California, (November, 1998).

19. T. W. Berger, M. Baudry, R. D. Brinton, B. Sheu, and A. R. Tanguay, Jr., "A Hybrid Neuron-Silicon Computational System for Pattern Recognition", DARPA Controlled Biological Systems Program Annual Program Review, Tucson, Arizona, (February, 1999).

20. M. Baudry, T. W. Berger, R. D. Brinton, B. Sheu, and A. R. Tanguay, Jr., "Hybrid Biological-Electronic Biosensor for Detection of Chemically- or Biologically-Induced Cognitive Impairment", DARPA Tissue Based Biosensors Program Annual Program Review, Tucson, Arizona, (February, 1999).

21. A. R. Tanguay, Jr. and B. K. Jenkins, "Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures and Smart Camera Applications", presentation for Matsushita Corporation, Center for Neural Engineering, University of Southern California, Los Angeles, California, (February 18, 1999).

22. A. R. Tanguay, Jr., "Emerging Smart Camera Technologies: Toward an Adaptive Optoelectronic Eye", Winter Conference on Neural Plasticity, Workshop on Hardware Implementations of Neural Networks, St. Lucia, West Indies, (February 23-26, 1999).

23. B. K. Jenkins, "3-D Photonic Artificial-Neural Systems With Applications to Vision," Winter Conference on Neural Plasticity, Workshop on Hardware Implementations of Neural Networks, St. Lucia, West Indies (Feb. 23-26, 1999).

24. B. K. Jenkins, P. Nasiatka, and A. R. Tanguay, Jr., "Use of VCSEL Arrays in 3-D Photonic Multichip Modules", Joint Optoelectronics Program (JOP) User Review Seminar, San Francisco, California, (March 29-30, 1999); (Invited Paper).

25. I. Biederman, "Recognizing Depth-Rotated Objects: A Review of Recent Research and Theory", Workshop on Visual Object Recognition by Humans and Machines, Bad Homburg, Germany, (May, 1999); (Invited Presentation).

26. O. Painter, R. K. Lee, A. Yariv, A. Scherer, J. D. O'Brien, I. Kim, and P. D. Dapkus, "Two-Dimensional Photonic Bandgap Defect Laser", Conference on Lasers and Electro-Optics 1999 (CLEO 99), Baltimore, Maryland, (May 23-28, 1999); Postdeadline Paper CPD21.

27. O. Painter, R. K. Lee, A. Yariv, A. Scherer, J. D. O'Brien, I. Kim, and P. D. Dapkus, "Two-Dimensional Photonic Bandgap Defect Laser", CLEO Europe 99, Munich, Germany, (June 13-17, 1999).

28. B. K. Jenkins and A. R. Tanguay, Jr., "Applications of 3-D Photonic Multichip Modules and 2-D Incoherent/Coherent Source Arrays", presentation for Matsushita Corporation, Center for Neural Engineering, University of Southern California, Los Angeles, California, (July 16, 1999).

29. O. Painter, R. K. Lee, A. Yariv, A. Scherer, J. D. O'Brien, I. Kim, and P. D. Dapkus, "Two-Dimensional Photonic Bandgap Defect Laser", LEOS Summer Topical Meetings 1999, San Diego, California, (July 27-30, 1999).

30. A. R. Tanguay, Jr., "Emerging Smart Camera Technologies: Toward an Adaptive Optoelectronic Eye", Eastman Kodak Company, Systems Concept Center, Rochester, New York, (August 4, 1999); (Invited Presentation).

31. A. R. Tanguay, Jr. and B. Keith Jenkins, "Hybrid Electronic/Photonic Multichip Modules for Vision and Neural Prosthetic Applications", National Institute of Mental Health/Alfred E. Mann Institute-USC Conference on Replacement Parts for the Brain: Intracranial Implantation of Hardware Models of Neural Circuitry, Washington, D.C., (August 12-14, 1999); (Invited Presentation).

32. B. Keith Jenkins and A. R. Tanguay, Jr., "Photonic Artificial-Neural Adaptive Systems with Applications to Vision," National Institute of Mental Health/Alfred E. Mann Institute-USC Conference on Replacement Parts for the Brain: Intracranial Implantation of Hardware Models of Neural Circuitry, Washington, D.C., (August 12-14, 1999); (Invited Presentation).

33. Armand R. Tanguay, Jr., "Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures", ARO MURI Adaptive Optoelectronic Eye Research Review, Army Research Laboratory, Adelphi, Maryland, (August 23-24, 1999).
34. Christoph von der Malsburg, "Wavelet-Based Vision Algorithms", ARO MURI Adaptive Optoelectronic Eye Research Review, Army Research Laboratory, Adelphi, Maryland, (August 23-24, 1999).
35. Gary Holt, "Development of Mappable Algorithms for Edge Detection and Image Feature Conjunctions", ARO MURI Adaptive Optoelectronic Eye Research Review, Army Research Laboratory, Adelphi, Maryland, (August 23-24, 1999).
36. B. Keith Jenkins, "Mapping Vision Processes onto Hybrid Electronic/Photonic Multichip Module Architectures", ARO MURI Adaptive Optoelectronic Eye Research Review, Army Research Laboratory, Adelphi, Maryland, (August 23-24, 1999).
37. John O'Brien, "Hardware Components for Hybrid Electronic/Photonic Multichip Module Integration", ARO MURI Adaptive Optoelectronic Eye Research Review, Army Research Laboratory, Adelphi, Maryland, (August 23-24, 1999).
38. I. Biederman, "An Evaluation of 'View-Based' *vs.* Geon Structural Descriptions as Alternative Accounts of Visual Object Recognition", 2nd IEEE Workshop on Generic Object Recognition, Corfu, Greece, (September, 1999); (Invited Presentation).
39. I. Biederman, "Aiding Image Analysts through RSVP Training and Displays", Meeting on Neuroscience Inspired Target Recognition", The Neuroscience Institute, La Jolla, California, (September, 1999); (Invited Presentation).
40. R. H. Tsai, J. C. Tai, B. J. Sheu, A. R. Tanguay, Jr., and T. W. Berger, "Design of a Scalable and Programmable Hippocampal Neural Network Multichip Module", Society for Neuroscience Annual Meeting, Miami, Florida, (October 23-28, 1999).
41. A. Kosta and I. Biederman, "Does Variability in the Size of an Object's Parts Facilitate Recognition?", 7th Annual Workshop on Object Perception and Memory, Los Angeles, California, (November, 1999).
42. E. A. Vessel, M. C. Mangini, and I. Biederman, "Experts *vs.* Novices Performing Subordinate RSVP Identification", 7th Annual Workshop on Object Perception and Memory, Los Angeles, California, (November, 1999).

43. M. C. Mangini and I. Biederman, "Do Objects with Many Parts Incur Greater Attentional Costs than Objects with Few Parts?", 7th Annual Workshop on Object Perception and Memory, Los Angeles, California, (November, 1999).

44. R. Vogels, I. Biederman, and M. Bar, "Sensitivity of Macaque Temporal Neurons to Variations in Object Shading", Meetings of the Society for Neuroscience, Miami, Florida, (November, 1999).

45. O. Painter, R. K. Lee, A. Yariv, A. Scherer, J. D. O'Brien, I. Kim, and P. D. Dapkus, "Photonic Bandgap Defect Laser", IEEE Lasers and Electrooptics Society (IEEE/LEOS) 1999 Annual Meeting, San Francisco, California, (November 8-11, 1999).

46. A. R. Tanguay, Jr., M. Han, and P. Nasiatka, "Hybrid Biological/Electronic/Photonic Devices", Joint DARPA Controlled Biological Systems (CBS) and Tissue Based Biosensors (TBB) Annual Program Review, University of Southern California, Los Angeles, California, (December 8, 1999).

47. A. R. Tanguay, Jr., "Adaptive Optoelectronic Eyes: Hybrid Sensor/Processor Architectures", ARL-SEDD/ARL-ARO Integrated Imaging Workshop, Army Research Office, Research Triangle Park, North Carolina, (December 17, 1999); (Invited Presentation).

48. I. Biederman, "Human Face and Object Recognition in Vertebrates (Man and Macaque)", Workshop on Recognition of Visual Patterns and Landmarks by Insects, Delmenhorst, Germany, (March, 2000); (Invited Presentation).

49. T. W. Berger, M. Baudry, R. D. Brinton, J. Liaw, V. Marmarelis, B. Sheu, and A. R. Tanguay, Jr., "A Hybrid Neuron-Silicon Computational System for Pattern Recognition", DARPA Controlled Biological Systems Program Annual Program Review, San Antonio, Texas, (April, 2000).

50. M. Baudry, T. W. Berger, R. D. Brinton, J. Liaw, V. Marmarelis, B. Sheu, and A. R. Tanguay, Jr., "Hybrid Biological-Electronic Biosensor for Detection of Chemically- or Biologically-Induced Cognitive Impairment", DARPA Tissue Based Biosensors Program Annual Program Review, San Antonio, Texas, (April, 2000).

51. I. Biederman, "Shape Recognition in Mind and Brain", Symposium on Object Recognition at the International Congress of Psychology, Stockholm, Sweden, (July, 2000); (Invited Presentation).

52. A. R. Tanguay, Jr., B. K. Jenkins, I. Biederman, C. von der Malsburg, B. Mel, J. O'Brien, and P. D. Dapkus, "Dense 3-D Integrated Photonic Multichip Modules for Adaptive Spatial and Spectral Image Processing Applications", DARPA Photonic Wavelength and Spatial Signal Processing (PWASSP) Kick-Off Meeting, Colonial Williamsburg, Virginia, (September 12-13, 2000).

53. A. R. Tanguay, Jr., B. K. Jenkins, I. Biederman, C. von der Malsburg, B. Mel, J. O'Brien, and P. D. Dapkus, "Dense 3-D Integrated Photonic Multichip Modules for Adaptive Spatial and Spectral Image Processing Applications", DARPA Microelectronics Technology Office Optoelectronics Annual Review, Cincinnati, Ohio, (October 16-19, 2000).
54. A. R. Tanguay, Jr., B. K. Jenkins, C. von der Malsburg, B. Mel, G. Holt, J. O'Brien, I. Biederman, A. Madhukar, P. Nasiatka, and Y. Huang, "Vertically Integrated Photonic Multichip Modules for Vision Applications", Symposium on Physical Optics for Digital Imaging/Digital Optics for Physical Imaging, David J. Brady, Symposium Organizer; Annual Meeting of the Optical Society of America, Providence, Rhode Island (October 22-27, 2000); (Invited Paper).
55. R. Vogels, I. Biederman, M. Bar, and A. Lorincz, "The Representation of Objects in Inferior Temporal Cortex (IT)", Annual Meeting of the Psychonomic Society, New Orleans, Louisiana, (November, 2000).
56. J. Zhu and C. von der Malsburg, "Fast Dynamic Link Matching", Fifth International Conference on Cognitive and Neural Systems (ICCNS), Boston University, (May 30 - June 2, 2001).
57. C. Kim, W. J. Kim, A. Stapleton, J. R. Cao, and J. O'Brien, "Quality Factors in Single Defect Photonic Crystal Lasers with Asymmetric Cladding Layers," Paper TuJ4, LEOS 2001, San Diego, CA, (2001).
58. W. J. Kim and J. D. O'Brien, "A Full Vectorial Analysis of 2-D Photonic Crystal Slabs," Paper TuJ3, LEOS 2001, San Diego, CA, (2001).
59. P.-T. Lee, J. R. Cao, S. J. Choi, Z.-J. Wei, J. D. O'Brien, and P. D. Dapkus, "Room Temperature Operation of VCSEL-Pumped Photonic Crystal Lasers," Paper ThA1, LEOS 2001, San Diego, CA (2001).
60. B. W. Mel, E. T. Ortega, C. Zhou, and G. R. Holt, "Seeing the Cartoon in a Complex Scene: Lessons from Visual Cortex", [Abstract], Society for Neuroscience Annual Meeting, San Diego, CA, #286.11, (2001).
61. J. Zhu and C. von der Malsburg, "Learning Control Units", Sixth International Conference on Cognitive and Neural Systems (ICCNS), Boston University, (May 30 - June 1, 2002).
62. J. R. Cao, P.-T. Lee, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Lithographic Tuning of 2-D Photonic Crystal Lasers," Paper TuW4, OFC 2002, Anaheim, CA, (2002).
63. J. O'Brien, P.-T. Lee, J.-R. Cao, C. Kim, W. J. Kim, S.-J. Choi, and P. D. Dapkus "VCSEL-Pumped Photonic Crystal Lasers," Paper CtuW3, CLEO 2002, Long Beach, CA, (2002); (Invited Presentation).

64. C. Zhou, and B. W. Mel, "Combining Multiple Cues for Contour Integration in Natural Scenes", [Abstract], Society for Neuroscience Annual Meeting, Orlando, FL, #260.21, (2002).

65. P.-T. Lee, J. R. Cao, S.-J. Choi, T. Yang, J. D. O'Brien, and P. D. Dapkus, "Investigation of the Optical Losses in Photonic Crystal Laser Cavities by Varying the Number of Lattice Periods," Paper ThB5, 2002 International Semiconductor Laser Conference, Garmisch-Partenkirchen, Germany, (2002).

66. J. O'Brien, "Photonic Crystal Lasers," SPIE Photonic Fabrication Europe Symposium, Brugge, Belgium; also SPIE Proceedings on VCSELs and Optical Interconnects, (2002); (Invited Presentation).

67. C. Zhou, and B. W. Mel, "A Probabilistic Approach to Cue Combination for Color Boundary Detection", [Abstract], Society for Neuroscience Annual Meeting, New Orleans, LA, #590.17, (2003).

68. J. O'Brien, J.-R. Cao, W. Kuang, M.-H. Shih, W. J. Kim, C. Kim, P.-T. Lee, S.-J. Choi, and P. D. Dapkus, "Photonic Crystal Devices," Paper JWC3, Optics in Computing 2003, Washington, D.C., (2003); (Invited Presentation).

69. J.-R. Cao, P. Lee, S. Choi, J. O'Brien, and P. D. Dapkus, "Threshold Pump Power Dependence on the Spectral Alignment Between the Gain Peak and the Cavity Resonance in InGaAsP Photonic Crystal Lasers," Paper MF57, OFC 2003, (2003).

70. J.-R. Cao, Z.-J. Wei, S.-J. Choi, W. Kuang, H. Yu, J. D. O'Brien, and P. D. Dapkus, "Sapphire Bonded Photonic Crystal Microcavity Lasers," Paper TuE5, Annual Meeting of the Optical Society of America, Tuscon, AZ, (2003).

71. P. D. Dapkus and J. D. O'Brien, "Meso- and Nanophotonic Devices for Integrated Photonic Circuits," Paper IV.5, Device Research Conference 2003, Salt Lake City, UT, (2003); (Invited Presentation).

72. J. D. O'Brien, "Photonic Crystal Devices," Paper WL1, LEOS Annual Meeting 2003, Tuscon, AZ, (2003); (Invited Presentation).

73. J. D. O'Brien, J.-R. Cao, W. Kuang, M. H. Shih, W. J. Kim, H. Yukawa, C. Kim, S. Choi, and P. D. Dapkus, "Photonic Crystal Waveguides and Emitters," Paper 5277-32, SPIE International Symposium on Microelectronics, MEMs, and Nanotechnology 2003, Perth, Australia, (2003); (Invited Presentation).

74. M. Huesken, M. Brauckmann, S. Gehlen, K. Okada and C. von der Malsburg, "Evaluation of Implicit 3D Modeling for Pose Invariant Face Recognition", SPIE Defense and Security Symposium, (April 12-16, 2004).

75. J. D. O'Brien, "Design, Fabrication, and Characterization of Photonic Crystal Waveguides," Paper 5359-30, Photonics West 2004, San Jose, CA, (2004); (Invited Presentation).

76. W. Kuang, J.-R. Cao, T. Yang, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Classification of Modes in Multi-Moded Photonic Crystal Microcavities," Paper CtuDD3, CLEO 2004, San Francisco, CA, (2004).
77. J.-R. Cao, W. Kuang, Z.-J. Wei, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Far-Fields of Photonic Crystal Microcavity Lasers," Paper CtuR7, CLEO 2004, San Francisco, CA, (2004).
78. W. Kuang and J. D. O'Brien, "Photonic Crystal Devices," The International Symposium on Optical Science and Technology: SPIE 49<sup>th</sup> Annual Meeting, Paper 5554-28, Denver, CO, August 3, 2004, (2004).
79. J. R. Cao, Z. Wei, S. Choi, W. Kuang, J. D. O'Brien, and P. D. Dapkus, "Sapphire-Bonded Photonic Crystal Lasers," TuA-3-2, 6th International Conference on Indium Phosphide and Related Materials (IPRM'04), Kagoshima, Japan, (2004).
80. J. R. Cao, W. Kuang, S.-J. Choi, J. D. O'Brien, and P. D. Dapkus, "Modified Photonic Crystal D<sub>3</sub> Laser Cavity for Improving Side Mode Suppression Ratio," Paper FB2, 19th IEEE International Semiconductor Laser Conference (ISLC'04), Matsue, Japan, (September, 2004).
81. J. D. O'Brien, J.-R. Cao, A. Stapleton, M.-H. Shih, W. Kuang, W. J. Kim, Z.-J. Wei, S.-J. Choi, and P. D. Dapkus, "Characterization of Photonic Crystal Structures," Session III, Symposium on Optical Fiber Measurements (SOFM 2004), Boulder, Colorado, (September 2004); (Invited Presentation).
82. J. D. O'Brien, J.-R. Cao, W. Kuang, M.-H. Shih, W. J. Kim, A. Stapleton, Z.-J. Wei, S.-J. Choi, and P. D. Dapkus, "Photonic Crystal Devices," Paper FMK5,, Frontiers in Optics/Laser Science XX Conference (88th Annual Meeting of the Optical Society of America), Rochester, New York, (October, 2004); (Invited Presentation).
83. J. D. O'Brien, "Nanophotonic Devices," International Workshop on Laser Cleaning 4, Sydney, Australia, (December 15, 2004); (Invited Presentation).
84. J. D. O'Brien, "Photonic Crystal Devices," COMMAD 04, Brisbane, Australia, (December, 2004); (Invited Presentation).
85. C. Zhou and B. W. Mel, "Combining Cues for Boundary Detection Using the "Mixture of Specialists" Model", [Abstract], COSYNE Conference, Salt Lake City, Utah, #295, (2005).



## **Report of Inventions**

No patent applications directly attributable to the MURI effort were disclosed, filed, or awarded during the research program.

## **Technology Transfer**

During the research program period, several of the investigators had significant interactions with the various DoD agencies, as well as with corporations and DoD contractors. In several cases, key research results that emerged during this research program are already in the early stages of technology transfer. These significant interactions and initial technology transfer efforts are an important program component, as described below.

Professor Irving Biederman developed an ongoing research collaboration with Dr. Barbara L. O'Kane of the U.S. Army CECOM RDEC Night Vision & Electronic Sensors Directorate, Ft. Belvoir, VA, on the identification of targets in infrared imagery by human observers. This research program looked at the full range of target images from where the parts and hot spots are well defined to when the vehicles look like dental fillings. A significant portion of this effort was directed toward the development of a system with ATR capabilities.

Professor Christoph von der Malsburg had extensive interactions with the Optical Tracking Group, Avionic Equipment Section, (Dr. Gabriel Udomkesmalee) at the Jet Propulsion Laboratory.

Prof. von der Malsburg also started a U.S. company (Eyematic Interfaces, Inc., Santa Monica, CA) for development of face recognition and tracking systems (Dr. Hartmut Neven, Director). Currently, visual recognition algorithms and software developed by Prof. von der Malsburg and his students are being employed by Eyematic for these applications.

Dr. von der Malsburg and his associates developed a face recognition system under contract from the Army Research Laboratory (ARL) under the FERET (Face Recognition Technology) Program (Dr. Jonathon Phillips, Contract Monitor). This system was repeatedly tested by ARL in competition to other groups, and has repeatedly outperformed the competing recognition systems. In one such test the USC group outperformed all other groups, including Dr. A. (Sandy) Pentland's group of the MIT Media Lab.

Under a second contract from the Army Research Laboratory (ARL), Prof. Von der Malsburg's group developed a Person Spotter System that is able to extract and recognize faces from live video streams.

Dr. von der Malsburg's research group also collaborated with Siemens Corporate Research, Munich (Prof. U. Ramacher) on the development of an advanced vision system (SEE-1). Its core is an array of digital signal processing chips (DSP's), and its design is optimized for the computing-intensive algorithms

developed in Dr. von der Malsburg's Laboratory. The SEE-1 will have a sustained processing power of 5 GOPS. This advanced vision system is intended for integration with the Person Spotter System, among other applications.

In a MURI-related project financed by the DoD Counterdrug Technology Development Program Office, and administered by the Army Research Laboratory, Dr. von der Malsburg's research group developed a system for the recognition of faces from live video. This system was developed in consultation with Mr. Tommy Walker, Naval Surface Warfare Center, Crane, Indiana.

In addition, a second MURI-related research project in Dr. von der Malsburg's research group, funded by the Office of Naval Research, was focused on the fusion of consecutive image frames for the purpose of improved target recognition and tracking.

Professor Armand R. Tanguay, Jr., and Professor B. Keith Jenkins developed a technical collaboration with the Army Research Laboratory (Dr. Joe Mait), on the design and application of novel subwavelength diffractive optical elements (DOE's). Professor Gregory D. Nordin, of the University of Alabama, Huntsville (UAH), an Affiliated Faculty Member of the MURI research program, was also involved in this collaborative effort. The interaction involved the development of new design and fabrication methods for novel DOE's, and combined key DOE design and analysis expertise from ARL, as well as extensive rigorous diffraction analysis expertise at UAH, with analytical and fabrication techniques that have evolved from the MURI research program at USC. This collaboration was also directed towards uncovering the fundamental and technological potential (as well as limitations) of adding such subwavelength capability to DOE's. The multi-group interaction resulted in the design and simulation of new elements at ARL, and in the evaluation of such subwavelength-feature DOE performance from the point of view of multilayer computational structures at USC. Potential applications include dense chip-to-chip optical interconnections as well as other diffractive optical systems.

Professor B. Keith Jenkins developed an interaction with technical personnel at the TRW Automotive Electronics and Space and Defense Divisions (Dr. Barry Dunbridge) on information display and driver interfaces in the automotive cockpit, which could potentially provide an applications vehicle for hybrid electronic/photonics computational modules, investigated within this MURI program.

Professor Armand R. Tanguay, Jr. transitioned key research results on single-sided flip-chip bonding technology to Teledyne Electronics Technologies (Marina Del Rey, CA; Mr. Robert Steenberge), a key corporate partner in the MURI effort. This technology may prove to be exceedingly useful in the packaging of multichip modules with industry-supplied OEM microprocessor, memory, DSP, and ASIC chip sets.

In a second interaction with a Teledyne company, Prof. Tanguay's group initiated a collaborative effort with Teledyne Lighting and Displays (Hawthorne,

CA; Dr. David Pelka) on the application of optical power bus technology in display backlighting configurations, as well as on the antireflection coating of microprismatic beam steering arrays.

Professor Tanguay developed an intensive interaction with the Eastman Kodak Company (Rochester, NY; Dr. Gary L. Bottger, Dr. John Spoonhower, Mr. Les Moore) on both immersive cameras and smart camera technology. The immersive camera concept provides a natural applications vehicle for adaptive optoelectronic eyes, and smart cameras can potentially provide key stepping stones along the way to a fully functional adaptive system implementation. A key current focus of the smart camera research project centers on contrast enhancement, color constancy, and chromatic differentiation for the disambiguation of camouflage (as well as for detection of smart fiducials in a natural environment). An additional feature of the project is the use of adaptive nonlinear dynamic range compression algorithms for the acquisition and processing of images in lighting conditions that span both bright (*e.g.*, sunlit) and dark (*e.g.*, shadowed) regions. Professor Tanguay spent six weeks of his sabbatical leave during each of the summers of 1999 and 2000 at Eastman Kodak Company, as well as at the University of Rochester, Institute of Optics (Prof. Nicholas George and Prof. Dennis Hall).

Professor John O'Brien undertook a collaboration with Agilent Laboratories on photonic crystal components for multi-wavelength processing, and as a result received additional financial support from Agilent for this program. As part of this research program, Agilent worked on developing an imprint lithography technique that has the potential for large scale production of these crystals. In addition, they have a beam writer that is capable of writing patterns over large areas. Agilent agreed to pattern photonic crystal waveguides for us over large areas to facilitate their characterization as part of the MURI research program as well.

Professor Anupam Madhukar had considerable interaction and cooperative work with the Avionics Division of WPAFB (Drs. Cole Litton and Edward Stutz) and the Electronics Division of the Army Laboratory at Ft. Monmouth (Drs. T. Aucoin, D. Smith, and P. Newman), including joint publications with the latter. He also had active interactions with the Avionics and Materials Divisions of WPAFB and developed interactions with the Army Research Laboratory (Dr. Richard Leavitt) in the context of IR detectors.

Furthermore, Prof. Madhukar was Principal Investigator of a related MURI effort that focused on IR detector arrays based on emerging quantum dot technology ("Stress-Engineered Quantum Dots for Multispectral Infrared Detector Arrays", FY 98 MURI Program, Contract No. F49620-98-1-0474; Program Manager: Maj. Daniel K. Johnstone, Air Force Office of Scientific Research). The goal of this related research program was to develop IR focal plane arrays with enhanced sensitivity and quantum efficiency by making use of the significant increase in absorption cross section that results from 2-D quantum confinement.

## Multidisciplinary Education

The impact of this MURI effort on USC's ability to conduct DoD-relevant research and educate students was significant in terms of both research integration and the development of human resources. Currently, most advanced sensor and electronic/photonic packaging research is carried out in industrial and government laboratories, and neither is currently established as a viable discipline. This MURI program integrated an unusually broad research team focused in both of these emerging technology areas, and greatly enhanced the university's capability to perform cutting edge research in these areas. This capability was further enhanced through the acquisition of novel capital equipment items funded by a separate MURI-related FY 99 DURIP equipment grant.

The MURI program attracted excellent prospective undergraduates and graduates. We produced a large number of highly interdisciplinary Ph.D. theses and B.S./M.S. theses during the course of the program, which will have a direct impact on the trained workforce available to industrial and government laboratories, as well as other universities.

During the grant period, several excellent new Ph.D. students were recruited for participation in the MURI research program, and overall thirty-four (34) Ph.D. students (partially funded or funded by related efforts) were involved directly in MURI-related research. Thirty (30) M.S., Engr., and Ph.D. degrees were granted to students who were involved (either directly or indirectly) with the MURI program, as listed above.

Extensive multidisciplinary interactions among all of the graduate and undergraduate research assistants were undertaken during this grant, including interactions with all of the MURI faculty members, through regularly scheduled technical meetings, held throughout the period at an average rate of eight per month: four on algorithm and architectural issues, two on hardware implementation issues, and two meetings of the entire MURI team on programmatic issues, integration of concepts, and cross-disciplinary fertilization (including reports from the two interacting groups).